2011

Biodiversity and emerging biogeography of the neutrophilic iron-oxidizing Zetaproteobacteria

Sean M. McAllister
Western Washington University

BIODIVERSITY AND EMERGING BIOGEOGRAPHY OF THE NEUTROPHILIC IRON-OXIDIZING *ZETAPROTEOBACTERIA*

By

Sean M. McAllister

Accepted in Partial Completion
Of the Requirements for the Degree
Master of Science

Moheb A. Ghali, Dean of the Graduate School

ADVISORY COMMITTEE

Chair, Dr. Craig Moyer

Dr. Dietmar Schwarz

Dr. Jeff Young

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Sean M. McAllister

June 10, 2011

BIODIVERSITY AND EMERGING BIOGEOGRAPHY OF THE NEUTROPHILIC IRON-OXIDIZING *ZETAPROTEOBACTERIA*

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Sean M. McAllister
June 2011

ABSTRACT

Members of the neutrophilic iron-oxidizing candidate class "*Zetaproteobacteria*" have predominantly been found at sites of microbially mediated iron oxidation in marine environments around the Pacific Ocean. Eighty-four full-length (>1,400 bp) and forty-eight partial-length *Zetaproteobacteria* small subunit ribosomal RNA (SSU rRNA) gene sequences from five novel clone libraries, one novel *Zetaproteobacteria* isolate, and the GenBank database were analyzed to assess the biodiversity of this burgeoning class of the *Proteobacteria* and to investigate its biogeography between three major sampling regions in the Pacific Ocean: Loihi Seamount, the Southern Mariana Trough, and the Tonga Arc. Sequences were grouped into operational taxonomic units (OTUs) based on a 97% minimum similarity. Of the 28 OTUs detected, 13 were found to be endemic to one of the three main sampling regions, and 2 were ubiquitous throughout the Pacific Ocean. Additionally, two deeply-rooted OTUs were identified that potentially dominate communities of iron-oxidizers originating in the deep subsurface. Spatial autocorrelation analysis and analysis of molecular variance (AMOVA) showed that geographic distance played a significant role in the distribution of *Zetaproteobacteria* biodiversity, whereas environmental parameters, such as temperature, pH, or total Fe concentration, did not have a significant effect. These results, detected using the coarse resolution of the SSU rRNA gene, indicate that the *Zetaproteobacteria* have a strong biogeographic signal.

ACKNOWLEDGEMENTS

I would like to thank the iron microbial observatory group (http://earthref.org/FEMO/), the operation teams for *Pisces V* and *Jason II*, and the captains and crew of the R/Vs *Kaimikai-o-Kanaloa*, *Melville*, *Kilo Moana*, and *Thomas G. Thompson* for their assistance with sample collection. I also thank Dietmar Schwarz and Benjamin Miner for their input toward this project's completion and invaluable assistance with statistical analyses. I extend my appreciation to the NSF-funded REU students working in Craig Moyer's lab for their aid in clone library construction and sequence data quality control, especially Travis Carney, Kelsey Leal, Sarah Safran, and Kyle Hager. Finally, I thank my committee members for their thoughtful comments on my draft theses. This project was funded in part by Western Washington University's Office of Research and Sponsored Programs and by the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The goal of biogeography is to study the distribution of an organism's biodiversity over space and time (43). With their small size, and thus great potential for global dispersal, there has been much debate over whether or not microorganisms can exhibit biogeography at all. Perhaps the most noted declaration/hypothesis was made by the Dutch microbiologist, Baas Becking, who said, "everything is everywhere: but the environment selects" (3, 51). Though perhaps highly simplified, this statement offers a good null hypothesis for microbial biogeography: that only the modern effects of environmental parameters play a significant role in the current distribution of microorganisms, not historical events such as dispersal or past habitat characteristics. Although some studies have supported this hypothesis (23, 41), many studies have been able to detect a nonrandom distribution of the microbe under investigation (e.g., reference 5), with some also showing geographically significant distribution patterns with little correspondence to observed environmental parameters (52, 70). Ultimately, the ability to detect the presence of extant biogeography has been shown to be dependent upon the spatial scale studied and the resolution of the selected molecular method (5, 31, 52).

In the deep ocean, sites of hydrothermal venting support a highly productive array of macrofaunal (60) and microbiological diversity driven by chemosynthesis (7, 20, 65). Associated with a large number of widely dispersed seamount, island arc, and ridge systems, hydrothermal vents are oases of life in the ocean, and as such are ideal systems with which to study biogeography. At seamounts, luxuriant Fe-rich microbial mats between 0.5 cm and 1 m thick are often observed, formed by iron-oxidizing bacteria (FeOB) oxidizing ferrous (Fe^{2+}) to ferric (Fe^{3+}) iron while fixing carbon (16, 19, 20). Diverse microorganisms, including

FeOB, have also been shown to thrive in fluids and sediments associated with hydrothermal systems (29, 35, 64). With over 125,000 seamounts worldwide (68), in addition to Fe-rich mats at backarc spreading centers (7) and mid-ocean ridge systems (20), deep-sea FeOB have the potential to play a considerable role in global Fe and carbon cycling, in addition to providing insight into the biogeography of hydrothermal vent-associated microbial communities. However, little is known about the formation and maintenance of these FeOB-dominated mat or fluid communities, or the global distribution and interaction of the dominant members of these communities.

It was initially assumed that the role of microbially-mediated iron-oxidation in the ocean (deep-sea or otherwise) was limited (19). This was assumed, in part, because the abiotic oxidation of Fe^{2+} to Fe^{3+} proceeds rapidly in oxygenated environments, in addition to the fact that Fe-oxidation produces minimal amounts of energy for growth; the current estimate for the energetic yield (ΔG°) from Fe^{2+} oxidation *in situ* is -90 kJ mol^{-1} of Fe^{2+} (16). Despite these perceived energetic limitations, FeOB have been found as dominant members of a large number of diverse environments, including freshwater systems (18, 61), deep-sea sediments (14), and sites of deep-sea hydrothermal venting associated with hotspot volcanism, island arc, and ridge systems (e.g., references 7, 13, 24, 34, 35, 55). At hydrothermal vents, with a large flux of Fe^{2+} estimated at 3×10^{11} mol per year and the production of steep redox gradients at the interface of vent effluent and cold seawater, these FeOB communities can thrive (25, 30).

The most common microscopic evidence for the activity of FeOB at sites of hydrothermal venting are tubular sheaths, helical stalks, y-shaped irregular filaments, and amorphous particles, all composed of Fe-oxyhydroxide excreted by the cell to avoid

encrustation as Fe-oxidation occurs (4, 19, 21, 36, 64). These structures have been found in both modern and ancient hydrothermal systems (33). Despite their abundance, FeOB have been historically difficult to culture. As a result, many of these structures were thought to belong to the fresh-water *Gallionella* spp. or *Leptothrix ochracea* (both *Betaproteobacteria*), which produce similar structures (16). However, to date, only one instance of a *Gallionella* phylotype has ever been reported at an active hydrothermal vent (34) (1 clone out of 127 in the library) and no *Leptothrix ochracea* have yet been detected. The question then is: what is oxidizing iron at hydrothermal vents? With the isolation of *Mariprofundus ferrooxydans* from the Fe-oxide dominant hydrothermal vents at Loihi Seamount and subsequent culture-independent molecular studies discussed herein, the *Zetaproteobacteria* have been identified as a diverse and abundant member of this deep-sea FeOB community (e.g., references 19, 21, 24, 28, 55).

Mariprofundus ferrooxydans is a chemolithoautotrophic, microaerophilic iron-oxidizing bacterium that grows in culture preferentially at 10-30˚C and circumneutral pH. Both strains of *M. ferrooxydans*, JV-1 and PV-1, produce filamentous stalk-like structures composed of Fe-oxyhydroxide. *M. ferrooxydans* is the only described representative of the *Zetaproteobacteria*, a novel, monophyletic candidate class of the *Proteobacteria* (21). The *Zetaproteobacteria* were first detected via culture-independent techniques by Moyer *et al.* (48) at Loihi Seamount, Hawaii. In that study, a single clone (PVB OTU4) was detected from a vent-associated microbial mat dominated by *Epsilonproteobacteria*.

Since this initial discovery, *Zetaproteobacteria* have been detected at several locations in diverse habitats around the world, including microbial mats and altered Fe-oxide-stained basalts at Loihi Seamount (13, 19, 21, 48, 56), microbial mats at the Southern

Mariana Trough (7), the brine-seawater interface at Kebrit Deep, Red Sea (12), microbial mat and basalt samples from Vailulu'u Seamount (64), mild steel corrosion enrichment experiments conducted in near-shore marine environments, Maine (44), and Fe-flocculent mats and sediments along the Kermadec Arc (29). However, the *Zetaproteobacteria* were not dominant members of the bacterial community in any of these studies. More recently, several studies focusing on low-temperature hydrothermal vent-associated microbial mats, sediments, and borehole fluids have shown the *Zetaproteobacteria* to be dominant and active members of these FeOB communities. These include studies from Loihi Seamount (55), off-axis Cleft Segment, Juan de Fuca Ridge (9), Tonga Arc (24), the Southern Mariana Trough (34, 35), and the Santorini flooded caldera, Greece (28). Even though the *Zetaproteobacteria* were initially thought to be rare, these studies have revealed 29 sites in 12 regions of the globe (predominantly in the Pacific Ocean) numbering more than 425 clones representing the *Zetaproteobacteria*. Of these, the vast majority (~73%) have been detected at seamounts. These studies have helped to focus our attention on these low-temperature seamount hydrothermal habitats, which seem to be where the *Zetaproteobacteria* are dominant.

Before we can test for the presence of biogeographical patterns in the *Zetaproteobacteria*, we must understand the currently sampled biodiversity, which has not yet been addressed. At present, with exception of the cultured isolates of *M. ferrooxydans*, this biodiversity has been sampled only at the level of the small subunit ribosomal RNA (SSU rRNA) gene. Our goal herein is to use these data, along with new SSU rRNA gene clone library data targeted to increase the sample size at Loihi Seamount, to describe *Zetaproteobacteria* operational taxonomic units (OTUs) (58). The increased sampling will allow us to assess the distribution of this biodiversity, therefore the biogeography, across

three major sampling regions: Loihi Seamount, the Southern Mariana Trough, and the Tonga Arc, each approximately equidistant (~6,000 km apart) in the Pacific Ocean. In addition, the clone libraries for this study were constructed from samples from multiple vent sites with readily available *in situ* chemistry data (25, 69). These novel sequences, when combined with sequence data from GenBank, will allow us to address the impact of environmental parameters on the global distribution and abundance of the *Zetaproteobacteria*. However, as found in previous studies, it is important to note that using the SSU rRNA gene for the study of biogeography offers only limited resolution (5, 31, 52). For this reason, further cultured isolates of the dominant members of the *Zetaproteobacteria* are needed.

This study is a primer for the investigation of *Zetaproteobacteria* biogeography. Further culturing efforts and studies focusing on the distribution patterns of the dominant *Zetaproteobacteria* OTUs identified herein will be necessary to identify small scale patterns of distribution that may exist between populations within major sampling regions of this deep-sea FeOB.

MATERIALS AND METHODS

Sample Collection

Five clone libraries were constructed from samples collected at Loihi Seamount, Hawaii, from 2004-2008 (Fig. S1, selected samples). Samples PV-601_b18 and PV-602_b14 were collected by suction sampler using *Pisces V* in 2004 (Upper Hiolo and Spillway sites, respectively). Samples J2-308_redgreen and J2-310_bluered were collected by suction sampler using *Jason II* in 2007 (Upper North Hiolo and Upper Lohiau sites, respectively). The J2-373_scoop1 clone library was constructed from a sample collected by scoop sampler using *Jason II* in 2008 (Pohaku site). After collection, all samples were stored at -80˚C until DNA extraction.

Genomic DNA Extraction

Genomic DNA (gDNA) was extracted from samples using the Fast DNA SPIN Kit for Soil (Qbiogene, Carlsbad, CA) according to the manufacturer's protocol with the modification that gDNA was eluted into 10 mM Tris with 0.1 mM EDTA at pH 8 (TE). To optimize the cellular lysis step, a FastPrep Instrument (Qbiogene) was used at an indexed speed of 5.5 for 30 sec. The purity and concentration of gDNA were determined with a NanoDrop ND-1000 spectrophotometer. All gDNAs were then diluted to ~10 ng/µl using TE buffer.

SSU rRNA Gene PCR Amplification and Clone Library Construction

Bacterial SSU rRNA genes were amplified from the gDNA using the 68F forward primer (5' TdNA dNAC ATG CAA GTC GdKdK CG 3') and the 1492R reverse primer (5' dKGdP TAC CTT GTT ACG ACT T 3'), where dK is a purine analog, dP is a pyrimidine analog, and dN is an equal mixture of dK and dP (Glen Research, Sterling, VA). Five

replicate PCRs were performed using 25-50 ng of gDNA template, 5 U of AmpliTaq Gold (Applied Biosystems, Carlsbad, CA), 1X AmpliTaq Gold PCR buffer, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 10 µg BSA, 1 µg T4g32p (Ambion, Austin, TX), 1 µM each of forward and reverse primers, and molecular grade water to a total volume of 50 µl. The following conditions were used for the amplification process: an initial 8-min hot-start at 95˚C, followed by 25-30 cycles of denaturation (94˚C for 1 min), annealing (58˚C for 90 sec), and elongation (72˚C for 3 min). This was followed by a final elongation step at 72˚C for 7 min. Amplicons were sized by 1% agarose gel electrophoresis against a 1-kb ladder (Invitrogen, Carlsbad, CA). Negative controls were maintained throughout. The five replicate PCRs were pooled, concentrated, and desalted with a Montáge PCR centrifugal filtration device (Millipore, Bedford, MA). The desalted PCR amplicons were then cloned with a TA cloning kit following the manufacturer's instructions (Invitrogen, Carlsbad, CA). All putative clones were streaked for isolation and the inserts assayed for correct size using PCR with M13F and M13R primers (46). Again, amplicons were sized against a 1-kb ladder using 1% agarose gel electrophoresis.

Plasmids were isolated and purified using standard alkaline lysis and then sequenced on an ABI 3130xl genetic analyzer. Initial operational taxonomic unit (OTU) composition for each clone library was determined based on reads from the 5' end of the SSU rRNA gene, and from one to three clones from each OTU were randomly selected for full-length sequencing using internal sequencing primers (38). SSU rRNA gene sequences were contiguously assembled (minimum 2X coverage) using BioNumerics (Applied Maths, Sint-Martens-Latem, Belgium).

Mariprofundus sp. strain M34 Isolation and Sequencing

Mariprofundus sp. strain M34 was isolated from sample J2-245_blue, which was collected by suction sampler using *Jason II* in 2006 (Spillway site). Freshly collected microbial mat was diluted directly into petri plates containing artificial seawater medium (ASW) with 1 μl/ml each of vitamins and mineral solutions (ATCC), and FeS as the iron source (17). Plates were incubated at room temperature in a sealed container with a BBL Campypak Plus microaerophilic system envelope (Becton, Dickinson and Co., Franklin Lakes, NJ). For the original enrichment, growth of stalk-forming, putative FeOB was observed by phase-contrast light microscopy in the 10^{-5} plate. Onshore, this enrichment was subjected to three more transfers of serial dilution to extinction. Each time the highest dilution that showed growth, typically 10^{-7}, was used for the subsequent transfer. Once a uniform cell/stalk morphology was observed that gave a consistent and unambiguous SSU rRNA gene sequence, the culture was checked for the presence of heterotrophic contaminants by streaking a sample on ASW-R2A agar plates. To confirm Fe lithotrophy, growth curves were completed with and without Fe present, which confirmed Fe^{2+} was required for growth (data not shown). Furthermore, the ability of the strain to grow in a liquid medium with $FeCl_2$ was confirmed to ensure that the strain was not growing on either sulfide or H_2 (17). A Mo Bio PowerSoil kit (Carlsbad, CA) was used to extract gDNA from the pure culture. The universal primers 27F (38) and 1492R (67) were used to amplify the SSU rRNA gene, with additional internal sequencing primers for full-length sequencing, as described above.

Zetaproteobacteria Sequence Recovery from GenBank and Chimera Screening

Published *Zetaproteobacteria* sequences were identified via two methods: NCBI's BLAST and the Ribosomal Database Project (RDP) Version 10.14 seqmatch algorithm (6).

All *Zetaproteobacteria* sequences were checked for chimeras using the Bellerophon server (32), RDP Version 8.1 chimera check v.2.7 (6), Pintail (*M. ferrooxydans* PV-1 used as reference sequence) (1), and Mallard (both *E. coli* and *M. ferrooxydans* PV-1 used as reference sequences) (2). No chimeras were detected among the full-length sequences used in this study. Two chimeras were identified in the partial-length sequence dataset (AB329957 and AB329967). Neither of these clones were used in this study.

OTU Determination

Sequences were categorized into full-length (>1,400 bp) only and full- plus partial-length datasets. The full-length dataset was trimmed to include data between the 68F and 1492R primers. The full- plus partial-length dataset was trimmed to include data between the universal priming sites 515F and 1406R. Priming sites were excluded in both datasets. These two datasets were aligned independently to the Arb-SILVA database using the SINA Webaligner function (53). Sequences were then masked so that phylogenetic/taxonomic analyses could be restricted to unambiguously aligned nucleotide positions.

Clones were then grouped into OTUs based on a minimum similarity of 97% (58). This similarity cutoff value has been widely accepted as the closest approximation for a microbial "species" short of culture-dependent analyses (59). OTUs were ranked based on the number of representative clones (e.g., OTU 1 contained the most clones and was thus the dominant OTU detected through this process).

Phylogenetic Analysis

Using the unambiguously aligned sequence data, phylogenetic placements according to maximum likelihood methods were calculated using fastDNAml version 1.2.2 (50) using the general two-parameter model of evolution (37) and allowing for the global swapping of

branches. The search for the optimal tree was repeated with these parameters until the best log likelihood tree was calculated in at least three independent tree calculations. The best tree for the full-length dataset was then bootstrapped 100 times allowing for global branch swapping. Due to computational constraints, the best tree for the full- plus partial-length dataset was bootstrapped 500 times without global branch swapping. For both datasets, the search for each bootstrap was repeated until the best log likelihood score was calculated for at least two independent bootstrap calculations.

Analysis of Molecular Variance (AMOVA)

Analysis of molecular variance (AMOVA) was conducted using Arlequin version 3.1.1 (22, 57). Sequences were organized by clone library and were then grouped by region, temperature, and sample type. Regional groupings were tested treating the southern Pacific Ocean both as one region and as three separate regions (e.g., Vailulu'u Seamount, Tonga Arc/East Lau Spreading Center, and Kermadec Arc). For temperature groupings, sequences were grouped by the temperatures of the environments from which they were isolated (psychrophilic [0-10˚C], mesophilic [11-40˚C], and [hyper]thermophilic [42-165˚C]). Where known, sequences were grouped by total Fe, Mn, and Si concentrations, Fe/Mn molar ratio, and pH, in addition to being grouped by region and temperature for these smaller datasets. AMOVA was also run separately with sequences belonging to OTUs 1 and 2, grouped by region and temperature. AMOVA was not done on the other OTUs due to limited sampling size. To test the affect of regional sample size on AMOVA results, a smaller subset of the database with only those sequences from the three main sampling regions (Loihi Seamount, the Southern Mariana Trough, and the southern Pacific Ocean group) was also run for all previously mentioned groupings. Full-length sequences were used for all tests, except for

when grouping by sample type (microbial mat, borehole fluid, and other), where both the full-length and full- plus partial-length datasets were used. The P-value significance tests for the variance components were carried out using 10,100 permutations.

Spatial Autocorrelation Analysis

Using the vegan package (version 1.17-2) of the R statistical analysis software environment (version 2.11.1), multivariate Mantel test statistics (r_M) were calculated to test for the presence of spatial autocorrelation (40, 49, 62). Euclidean geographic distances were calculated between sample sites using the reported geographic coordinates for published sequences in addition to geographic coordinates provided by ROV navigational data. These data were organized into a simplified geographic distance matrix where distances were broken into *d* classes with equal frequency of pairwise comparisons between classes. The similarity matrix for genetic distance between sample sites was calculated using the abundance-weighted non-normalized UniFrac distance metric (Fast UniFrac) (27, 42). The computed Mantel test statistic was tested for significance at $\alpha = 0.05$ using 999 permutations. Significance was determined from probability values corrected using the Bonferroni (conservative) and Holm methods. A Mantel correlogram (40) was created by plotting the Mantel test statistic against the previously determined distance classes. Only those sample sites with full-length sequences representing four or more clones from the three main sampling regions were used in this analysis.

Nucleotide Sequence Accession Numbers

The novel SSU rRNA gene sequences from this study have been submitted to GenBank and assigned accession numbers JF317957 (for *Mariprofundus* sp. strain M34) and JF320713 through JF320787 (for sequences listed in Table S1).

RESULTS

Clone Library and GenBank Recovery

Results for clone libraries constructed for this study from Loihi Seamount are summarized in Table S1. Most samples clustered into the broad Loihi Group I (dominated by members of the *Zetaproteobacteria*, *Gammaproteobacteria*, *Nitrospira*, and *Chloroflexi*) and Loihi Group II (dominated by members of the *Epsilonproteobacteria* and *Nitrospira*) categories as previously discussed (8, 20). Clone libraries PV-602_b14 (SPL) and J2-373_scoop1 (Poh) clustered as Loihi Group I, both dominated by the *Zetaproteobacteria*. Clone libraries PV-601_b18 (UHO) and J2-308_redgreen (UNH) clustered into Loihi Group II, dominated by *Nitrospira/Epsilonproteobacteria* and *Nitrospira*, respectively. Clone library J2-310_bluered (ULoh), dominated by *Actinobacteria* and *Deltaproteobacteria*, did not cluster into either broad category. In total, out of 74 full-length sequenced clones, 27 sequences belonged to the *Zetaproteobacteria*.

After collecting additional *Zetaproteobacteria* sequences from GenBank and screening all sequences for chimeras, the full-length sequence dataset consisted of 84 sequences masked to 1282 bp of unambiguously aligned positions and the full- plus partial-length dataset consisted of 132 sequences masked to 696 bp of unambiguously aligned positions. The majority of these sequences came from sites of hydrothermal venting around the Pacific Ocean (Fig. 1). Clone library and cultured isolate information is summarized in Table 1. *Zetaproteobacteria* were detected from a variety of habitats, including microbial mats, sediments, and borehole fluids, from psychrophilic (1.7˚C) to hyperthermophilic (165˚C) temperatures, with an average temperature of 32˚C. Approximately half of these clone libraries contained more than 10% *Zetaproteobacteria* clones.

OTU Designations

In the full-length sequence dataset, 28 OTUs were detected. With the addition of 48 sequences in the full- plus partial-length dataset, only 6 additional OTUs were detected and OTU designations did not show a large amount of variability from those of the full-length dataset (data not shown). With the smaller mask for this dataset (696 bp versus 1282 bp) leading to the omission of three out of six variable regions found in association with *Zetaproteobacteria* SSU rRNA secondary structures (Fig. S2), this partial-length dataset was not used in the statistical analyses, except where noted. However, due to the limited sample size of full-length sequences from the Southern Mariana Trough, both full- and partial-length sequences were used in regional comparisons.

A summary of the OTU designations for the full-length dataset can be seen in Table S2. Of the *Zetaproteobacteria* phylotypes detected there were 17 OTUs with three or more representative clones (OTUs 1-17), 11 OTUs containing at least three clones from more than one vent site (OTUs 1-4, 6, 8-11, 14, and 16), and 8 OTUs containing at least three clones from more than one geographic region (OTUs 1-4, 8, 9, 11, and 14). Partial-length sequences from the Southern Mariana Trough consisted of a number of clones grouping in OTU 1 (n=32), OTU 9 (n=28), and OTU 15 (n=14). The full-length sequences that made up the top 11 OTUs, in addition to OTU 15, are identified in the maximum likelihood tree (Fig. 2). Three of the four cultured *Zetaproteobacteria* isolates (including *M. ferrooxydans* strain PV-1, *M. ferrooxydans* strain JV-1, and *Mariprofundus* sp. strain M34) grouped into the eleventh most abundant OTU, which also included two environmental isolates, one from Loihi Seamount (Loh OTU7 clone 5) and the other from San Francisco Bay (WSMO200).

When considering the full- plus partial-length dataset, six major OTUs were found to comprise 72% of the *Zetaproteobacteria* diversity (OTUs 1-4, 9, and 15; Fig. 3). OTU 1 consisted of 67 full-length clones from Loihi Seamount (Marker [Mkr] #s 2-5, 39, 55, and 57), the Southern Mariana Trough (Fryer Site), the Tonga Arc (Volcanoes 1 and 19), the East Lau Spreading Center (TVG9), and the Kermadec Arc (Tangaroa Floc), with an additional 32 partial-length clones from the Southern Mariana Trough (Fryer and Kaiko Sites) and 3 partial-length clones from Loihi Seamount (Mkr #48). OTU 2 consisted of 54 full-length clones from Loihi Seamount (Mkr #s 2-5, 34, 36, 39, 55, 57, and Pele's Vents), the Juan de Fuca Ridge (off-axis Cleft Segment), and the Vailulu'u Seamount (Nafanua summit), with 6 additional partial-length clones from the Southern Mariana Trough (Fryer Site). The first environmental clone of the *Zetaproteobacteria*, PVB OTU4, was found to belong to this OTU. OTU 3 consisted of 36 full-length clones from Loihi Seamount (Mkr #57), the Juan de Fuca Ridge (off-axis Cleft Segment), and the Tonga Arc (Volcanoes 1 and 19). OTU 4 consisted of 34 full-length clones from Loihi Seamount (Mkr #57), the Tonga Arc (Volcano 1), and the East Lau Spreading Center (TVG9), with 3 additional partial-length clones from Loihi Seamount (Mkr #48). OTU 9 consisted of 8 full-length clones from the Southern Mariana Trough (Pika Site), the Juan de Fuca Ridge (off-axis Cleft Segment), Vailulu'u Seamount (Nafanua summit), and Boothbay Harbor, Maine, with 28 additional partial-length clones from the Southern Mariana Trough (Fryer and Pika Sites). OTU 15 consisted of 3 full-length clones from the Southern Mariana Trough (Pika Site), with 14 additional partial-length clones also from the Southern Mariana Trough (Pika and Kaiko Sites). OTUs 9 and 15 were deeply-rooted in the full- plus partial-length *Zetaproteobacteria* maximum likelihood

tree (Fig. 3) and the majority of the sequences from these OTUs were from borehole fluids taken from the Southern Mariana Trough by Kato *et al*. (35).

Regional Comparisons

Analysis of the OTU distribution between the three major sampling regions revealed several interesting patterns in biogeography (Fig. 4). Two OTUs (OTUs 1 and 2) were found to be ubiquitous throughout the Pacific Ocean. OTU 1 was found to be consistently present as a dominant member at all three sampling sites (within the top two OTUs detected representing more than 20% of the *Zetaproteobacteria* clones per site). Although OTU 2 was detectable throughout the Pacific Ocean, it was only found to be dominant at Loihi Seamount (~33% of the clones in that region). Though each region shares these two ubiquitous OTUs, each region hosts a unique diversity of the remaining, less abundant 26 OTUs (Fig. 4). Endemic OTUs, those unique to each site, were found at all three major sampling regions, numbering between two and six OTUs per site (Fig. 5). Comparisons between mat and borehole fluid sample types at the Southern Mariana Trough provided evidence for both geographic and environmental impact on OTU distribution. Borehole fluid samples, presumably originating in the deep subsurface, showed a lower richness and were dominated by OTUs either absent or detected at low levels in the overlying mat samples (Fig. 4).

Analysis of Molecular Variance (AMOVA)

Over forty separate AMOVA runs were made by grouping sequences as laid out in the methods section (data from Tables 1 and S3). The results of pertinent AMOVA runs are summarized in Table 2. Grouping all sequences by region showed little difference whether considering the southern Pacific Ocean as a single group (8.44% among group variation) or as three separate groups (8.86% among group variation; data not shown). Both among group

variance components were found to be significantly different from zero ($P \leq 0.05$). Considering this, all values are reported with the southern Pacific Ocean as a single region. Grouping all sequences by biologically relevant temperature preferences (psychrophilic, mesophilic, and [hyper]thermophilic) did not explain a significant amount of variation. Similarly, when considering smaller sequence subsets with known associated concentrations of Fe, Mn, and Si, Fe to Mn molar ratios, and pH, among group variation was not significantly different from zero (data shown for sequences grouped by iron concentration only; Table 2). The only other factor besides regional differences that explained a significant amount of sequence variation was sample type at 6.28% and 15.81% among group variation for the full-length and full- plus partial-length datasets, respectively. AMOVA results for a smaller dataset consisting of only the three main sampling regions showed similar results as the AMOVA run with the entire sequence dataset (data not shown). Regional groupings of OTU 2, found predominantly at Loihi Seamount, explained 29.96% of the sequence variability, the largest among group variance component detected. Both single OTU datasets showed higher among group variance components for regional groupings as compared to the temperature groupings. However, neither region nor temperature groupings for either single OTU dataset had among group variance components that were significant, though the regional groupings were nearly significant. This was likely due to the limited sampling size. In an attempt to compensate for sampling size while still testing for regional and environmental differences in closely related OTUs, and considering that sample type was found to explain a significant amount of genetic variation, AMOVA was run on a subset with only samples collected from microbial mats. For this subset, regional groupings continued to explain a significant amount of variation (8.31%) as compared to the temperature groupings,

which were not significantly different from zero. Significant sequence variability was detected within clone libraries for those runs including all sequences (n=12), accounting for an average of 67% of the total variation.

Spatial Autocorrelation Analysis

Spatial autocorrelation analysis utilizes the multivariate Mantel test statistic to test the null hypothesis that geographic distance is not correlated to the distribution of genetic diversity between sample sites (40, 49, 62). The abundance-weighted non-normalized UniFrac distance metric was calculated to compare genetic distance between sites (Fig. S3). Euclidean geographic distances were divided into three classes: class 1 (0-1,500 km, 28 pairwise comparisons), class 2 (3,300-5,500 km, 28 pairwise comparisons), and class 3 (5,500-6,600 km, 22 pairwise comparisons). The null hypothesis was rejected for distance class 1, which showed significant positive spatial autocorrelation ($P = 0.009$ to 0.011; probabilities corrected using the Holm and Bonferroni methods, respectively). Distance classes 2 and 3 did not show significant spatial autocorrelation ($P = 0.856$ and 0.354, respectively; corrected using the Bonferroni method), though they indicated a trend toward negative spatial autocorrelation with increasing distance. Results of the spatial autocorrelation analysis were plotted as a Mantel correlogram (Fig. 6).

DISCUSSION

Our understanding of the diversity and distribution of the *Zetaproteobacteria* is only beginning to emerge, despite their common occurrence at an increasing variety of hydrothermal vent sites. However, with the construction of five new clone libraries from Loihi Seamount, and the analysis of *Zetaproteobacteria* biodiversity and biogeography with these and additional clones from GenBank from across the Pacific Ocean, we have been able to identify 28 *Zetaproteobacteria* OTUs, some of which were found to be ubiquitous throughout the Pacific Ocean while others were endemic to the regions from which they were detected. Endemic OTUs may be found to be more cosmopolitan across the three main regions with additional sampling resulting in nearly full-length SSU rRNA gene sequences. Although 28 OTUs were identified, it is important to note that some of these OTU groupings disagreed with the phylogenetic placements of clones in the maximum likelihood trees (Fig. 2 and 3). These discrepancies highlight the difference between taxonomic and phylogenetic approaches (54). In most cases, phylotype and OTU are synonymous, especially when talking about quite distinct organisms. However, in the case of this study on *Zetaproteobacteria*, where many of the sequences were quite similar, this was not always true. Even though the OTUs and "phylotypes" agreed most of the time with respect to phylogenetic trees, there were times when separate OTUs were defined from what would probably be considered a single phylotype (such as OTUs 1 and 3 or OTUs 4 and 5). Unfortunately, no standardized definition of a phylotype has yet arisen, making the OTU the next best tool available. With these discrepancies, however, 28 *Zetaproteobacteria* OTUs might be slightly overestimated. A more conservative estimate of diversity would be to look at all those OTUs containing three or more clones (these OTUs are also less likely to contain

chimeric sequences). In this dataset, 17 OTUs with three or more clones were identified. This is still a substantial amount of previously unrecognized *Zetaproteobacteria* biodiversity.

Previous studies have found the conserved nature of the SSU rRNA gene to limit the resolution of biogeographic studies (5, 31, 52). Even with this coarse resolution, however, we were able to detect a non-random distribution of *Zetaproteobacteria* clones over geographic distances of ~6,000 km, with no significant impact from the environmental parameters that were tested. Initial observations of OTU distributions between the three main sampling regions identified thirteen endemic OTUs with ten other OTUs that were only shared between two of the main regions. Further analyses, including AMOVA and spatial autocorrelation analysis, were conducted to test the statistical validity of these observations. AMOVA run on the full-length dataset found that regional groupings could explain a significant percent of the genetic variation, whereas groupings by environmental parameters were not found to be significantly different from zero. Significant positive spatial autocorrelation was detected between samples separated by the lowest geographic distance (0-1,500 km; distance class 1). This positive spatial autocorrelation indicates that it is more likely for similar phylotypes to be found at this distance class than other distance classes with larger sample site separation, pointing to a non-random geographic distribution (40). Considering these data, Baas Becking's null hypothesis for the global mixing of all microorganisms can be rejected for the *Zetaproteobacteria*. At least for those populations surveyed in the Pacific Ocean, biogeography exists and was detectible using the coarse resolution of the SSU rRNA gene. It is possible that this strong biogeographic signal may be a result of the dispersal rate limitation that island-like relatively isolated hydrothermal vents

may maintain (66). This study adds to a growing number that have found microorganisms to have a more complex distribution than originally anticipated (5, 43, 52, 70).

Two of the OTUs identified in this study, OTUs 9 and 15, were found to be deeply-rooted in the *Zetaproteobacteria* tree and were supported by relatively high bootstrap values. Although a few sequences from Vailulu'u Seamount, Juan de Fuca Ridge, and Maine could be found in OTU 9, the vast majority of the clones from these two OTUs (~87%) originated at depth from borehole fluids collected from the deep subsurface at the Southern Mariana Trough (35). AMOVA runs grouping clone libraries by sample type found that a significant percentage of variation was explained by these groupings, though this result may be influenced by covariance with regional groupings. However, when only the samples from the Southern Mariana Trough were considered, sample type continued to play a considerable role in explaining the phylogenetic groupings of the *Zetaproteobacteria* (Fig. 3 and 4). This observation of distinctive OTU composition and diversity between microbial mat and borehole fluid communities over multiple sampling sites in a region suggests that there may be a community of *Zetaproteobacteria* endemic to the deep subsurface. Even with renewed interest in the deep biosphere, many questions regarding colonization and how life from the deep subsurface might interact with life at the seafloor remain unanswered. The *Zetaproteobacteria*, with members found both at and below the seafloor, may provide insight into these questions, and future studies of the *Zetaproteobacteria* should include a focus on these deep-subsurface OTUs and their detection and investigation at other sites around the world.

With the detection of biogeography at the coarse resolution of the SSU rRNA gene, it is likely that even stronger spatial patterns could be observed with finer levels of resolution

utilizing whole-genome scale approaches (52, 54). Genomics and metagenomics will also allow us to explore the metabolic diversity of these FeOB, as well as the idea of the ecotype or community of microorganisms as the unit of microbial evolution and ecology (11, 15, 26). At present, the most reliable methods for genomic studies involve the isolation of the microbe under investigation. Currently there are four isolates of the *Zetaproteobacteria*: *Mariprofundus ferrooxydans* strain PV-1, *M. ferrooxydans* strain JV-1, *Mariprofundus* sp. strain M34, and *Mariprofundus* sp. strain GSB2. Unfortunately, none of these isolates represent the majority of the environmental clones that have been detected (these isolates grouped, at best, in the 11th most abundant OTU). Thus, an important outcome of this study is the identification of phylotypes that should be targeted for future isolation attempts. We have already identified six OTUs that made up nearly three-quarters of the *Zetaproteobacteria* biodiversity: OTUs 1, 2, 3, 4, 9, and 15. These OTUs represent the breadth of the known *Zetaproteobacteria* biodiversity, and include dominant members of seafloor and sub-surface FeOB communities. OTUs 1 and 2 were also found to be ubiquitous throughout the main sampling regions in the Pacific Ocean. With the observation that the majority of *Zetaproteobacteria* diversity has been detected at mesophilic temperatures (Table 1), an observation in agreement with previous studies (24, 55), isolation attempts should be directed toward lower temperature hydrothermal habitats. With an average 67% of genetic variability found within clone libraries, the richness of OTUs at any one sample site should aid in future attempts at isolation, though perhaps these communities of putative FeOB share a syntrophic relationship, another reason why isolation has been so difficult in the past.

The *Zetaproteobacteria*, though detected, have not been found to be dominant at every site discussed in this study. Hodges and Olson (29) and Sudek *et al.* (64) found only seven *Zetaproteobacteria* clones combined, even though abundant Fe-oxyhydroxide sheaths were present at both sites. These results suggest that we may not fully understand the ecology of iron-oxidizing bacterial communities. A few hypotheses have been suggested: 1) There may be iron-oxidizers at hydrothermal vents other than the *Zetaproteobacteria* (20). 2) The *Zetaproteobacteria* may only be active in rapidly accreting mats (29, 55). 3) The sheath structure may be a result of the nucleation of poorly ordered Fe-oxyhydroxides or the adsorption of pre-existing Fe-oxide structures onto the surfaces of microbial cells, and may not necessarily indicate that iron oxidation is occurring (39, 63). Further attempts at isolating these non-*Zetaproteobacteria* FeOB should also be made. In addition, it seems possible and perhaps even probable that not all *Zetaproteobacteria* are Fe-oxidizers (20). Morphological comparisons using molecular tools such as fluorescence *in situ* hybridization (FISH) to link stalk, sheath, and y-shaped filament structures with phylogeny, cultivation-dependent studies, and single-cell genomics are all techniques that may be able to help unravel some of these questions.

Currently, the *Zetaproteobacteria* are the only known Fe-oxidizers growing at deep-sea hydrothermal vents. Understanding these FeOB is important for understanding the cycling of Fe and carbon at hydrothermal vents and potentially other marine sedimentary environments. With only three major sampling regions, more clones and isolates from more dispersed sampling sites are still required to more fully recognize the diversity, biogeography, and metabolic potential of the *Zetaproteobacteria*.

REFERENCES

1. **Ashelford, K. E., N. A. Chuzhanova, J. C. Fry, A. J. Jones, and A. J. Weightman.** 2005. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Appl. Environ. Microbiol. **71:**7724-7736.

2. **Ashelford, K. E., N. A. Chuzhanova, J. C. Fry, A. J. Jones, and A. J. Weightman.** 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. Appl. Environ. Microbiol. **72:**5734-5741.

3. **Baas Becking, L. G. M.** 1934. Geobiologie of inleiding tot de milieukunde. Van Stockum & Zoon, The Hague, The Netherlands.

4. **Chan, C. S., S. C. Fakra, D. Emerson, E. J. Fleming, and K. J. Edwards.** November 2010, posting date. Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation. ISME J. doi:10.1038/ismej.2010.173.

5. **Cho, J.-C., and J. M. Tiedje.** 2000. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. Appl. Environ. Microbiol. **66:**5448-5456.

6. **Cole, J. R., Q. Wang, E. Cardenas, J. Fish, B. Chai, R. J. Farris, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, T. Marsh, G. M. Garrity, and J. M. Tiedje.** 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res. **37**(Database issue): D141-D145. doi:10.1093/nar/gkn879.

7. **Davis, R. E., and C. L. Moyer.** 2008. Extreme spatial and temporal variability of hydrothermal microbial mat communities along the Mariana Island Arc and southern Mariana back-arc system. J. Geophys. Res. **113:**B08S15. doi:10.1029/2007JB005413.

8. **Davis, R., C. Moyer, S. McAllister, A. Rassa, and B. Tebo.** 2010. Spatial and temporal variability of microbial communities from pre- and post-eruption microbial mats collected from Loihi Seamount, Hawaii. Abstr. 13th International Symposium on Microbial Ecology, abstr. PS.01.015.

9. **Davis, R. E., D. S. Stakes, C. G. Wheat, and C. L. Moyer.** 2009. Bacterial variability within an iron-silica-manganese-rich hydrothermal mound located off-axis at the Cleft Segment, Juan de Fuca Ridge. Geomicrobiol. J. **26:**570-580.

10. **Dhillon, A., A. Teske, J. Dillon, D. A. Stahl, and M. L. Sogin.** 2003. Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. Appl. Environ. Microbiol. **69:**2765-2772.

11. **Doolittle, W. F., and O. Zhaxybayeva.** 2010. Metagenomics and the units of biological organization. Bioscience **60:**102-112.

12. **Eder, W., L. L. Jahnke, M. Schmidt, and R. Huber.** 2001. Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. Appl. Environ. Microbiol. **67:**3077-3085.

13. **Edwards, K. J., B. T. Glazer, O. J. Rouxel, W. Bach, D. Emerson, R. E. Davis, B. M. Toner, C. S. Chan, B. M. Tebo, H. Staudigel, and C. L. Moyer.** May 2011, posting date. Ultra-diffuse hydrothermal venting supports Fe-oxidizing bacteria and massive umber deposition at 5000 m off Hawaii. ISME J. doi:10.1038/ismej.2011.48.

14. **Edwards, K. J., D. R. Rogers, C. O. Wirsen, and T. M. McCollom.** 2003. Isolation and characterization of novel psychrophilic, neutrophilic, Fe-oxidizing, chemolithoautotrophic α- and γ-*Proteobacteria* from the deep sea. Appl. Environ. Microbiol. **69:**2906-2913.

15. **Elsaied, H., H. W. Stokes, T. Nakamura, K. Kitamura, H. Fuse, and A. Maruyama.** 2007. Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents. Environ. Microbiol. **9:**2298-2312.

16. **Emerson, D., E. J. Fleming, and J. M. McBeth.** 2010. Iron-oxidizing bacteria: an environmental and genomic perspective. Annu. Rev. Microbiol. **64:**561-583.

17. **Emerson, D., and M. M. Floyd.** 2005. Enrichment and isolation of iron-oxidizing bacteria at neutral pH. Meth. Enzymol. **397:**112-123.

18. **Emerson, D., and C. Moyer.** 1997. Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH. Appl. Environ. Microbiol. **63:**4784-4792.

19. **Emerson, D., and C. L. Moyer.** 2002. Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition. Appl. Environ. Microbiol. **68:**3085-3093.

20. **Emerson, D., and C. L. Moyer.** 2010. Microbiology of seamounts: common patterns observed in community structure. Oceanography **23:**148-163.

21. **Emerson, D., J. A. Rentz, T. G. Lilburn, R. E. Davis, H. Aldrich, C. Chan, and C. L. Moyer.** 2007. A novel lineage of *Proteobacteria* involved in formation of marine Fe-oxidizing microbial mat communities. PLoS ONE **2:**e667. doi:10.1371/journal.pone.0000667.

22. **Excoffier, L., G. Laval, and S. Schneider.** 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol. Bioinform. Online **1:**47-50.

23. **Finlay, B. J.** 2002. Global dispersal of free-living microbial eukaryote species. Science **296:**1061-1063.

24. **Forget, N. L., S. A. Murdock, and S. K. Juniper.** 2010. Bacterial diversity in Fe-rich hydrothermal sediments at two South Tonga Arc submarine volcanoes. Geobiology **8:**417-432. doi:10.1111/j.1472-4669.2010.00247.x.

25. **Glazer, B. T., and O. J. Rouxel.** 2009. Redox speciation and distribution within diverse iron-dominated microbial habitats at Loihi Seamount. Geomicrobiol. J. **26:**606-622.

26. **Green, J. L., B. J. M. Bohannan, and R. J. Whitaker.** 2008. Microbial biogeography: from taxonomy to traits. Science **320:**1039-1043.

27. **Hamady, M., C. Lozupone, and R. Knight.** 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. ISME J. **4:**17-27.

28. **Handley, K. M., C. Boothman, R. A. Mills, R. D. Pancost, and J. R. Lloyd.** 2010. Functional diversity of bacteria in a ferruginous hydrothermal sediment. ISME J. **4:**1193-1205. doi:10.1038/ismej.2010.38.

29. **Hodges, T. W., and J. B. Olson.** 2009. Molecular comparison of bacterial communities within iron-containing flocculent mats associated with submarine volcanoes along the Kermadec Arc. Appl. Environ. Microbiol. **75:**1650-1657.

30. **Holland, H. D.** 2006. The oxygenation of the atmosphere and oceans. Phil. Trans. R. Soc. B **361:**903-915.

31. **Huber, J. A., D. A. Butterfield, and J. A. Baross.** 2006. Diversity and distribution of subseafloor Thermococcales populations in diffuse hydrothermal vents at an active deep-sea volcano in the northeast Pacific Ocean. J. Geophys. Res. **111:**G04016. doi:10.1029/2005JG000097.

32. **Huber, T., G. Faulkner, and P. Hugenholtz.** 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics **20:**2317-2319.

33. **Juniper, S. K., and Y. Fouquet.** 1988. Filamentous iron-silica deposits from modern and ancient hydrothermal sites. Can. Mineral. **26:**859-869.

34. **Kato, S., C. Kobayashi, T. Kakegawa, and A. Yamagishi.** 2009a. Microbial communities in iron-silica-rich microbial mats at deep-sea hydrothermal fields of the Southern Mariana Trough. Environ. Microbiol. **11:**2094-2111.

35. **Kato, S., K. Yanagawa, M. Sunamura, Y. Takano, J. Ishibashi, T. Kakegawa, M. Utsumi, T. Yamanaka, T. Toki, T. Noguchi, K. Kobayashi, A. Moroi, H. Kimura, Y. Kawarabayasi, K. Marumo, T. Urabe, and A. Yamagishi.** 2009b. Abundance of *Zetaproteobacteria* within crustal fluids in back-arc hydrothermal fields of the Southern Mariana Trough. Environ. Microbiol. **11:**3210-3222.

36. **Kennedy, C. B., S. D. Scott, and F. G. Ferris.** 2003. Ultrastructure and potential sub-seafloor evidence of bacteriogenic iron oxides from Axial Volcano, Juan de Fuca Ridge, north-east Pacific Ocean. FEMS Microbiol. Ecol. **43:**247-254.

37. **Kishino, H., and M. Hasegawa.** 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. **29:**170-179.

38. **Lane, D. J.** 1991. 16S/23S rRNA sequencing, p. 115-175. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons Ltd., Chichester, UK.

39. **Langley, S., P. Igric, Y. Takahashi, Y. Sakai, D. Fortin, M. D. Hannington, and U. Schwarz-Schampera.** 2009. Preliminary characterization and biological reduction of putative biogenic iron oxides (BIOS) from the Tonga-Kermadec Arc, southwest Pacific Ocean. Geobiology **7:**35-49.

40. **Legendre, P., and L. Legendre.** 1998. Numerical ecology, 2nd English ed. Elsevier Science B.V., Amsterdam, The Netherlands.

41. **Lepage, E., E. Marguet, C. Geslin, O. Matte-Tailliez, W. Zillig, P. Forterre, and P. Tailliez.** 2004. Molecular diversity of new *Thermococcales* isolates from a single area of hydrothermal deep-sea vents as revealed by randomly amplified polymorphic DNA fingerprinting and 16S rRNA gene sequence analysis. Appl. Environ. Microbiol. **70:**1277-1286.

42. **Lozupone, C., and R. Knight.** 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. **71:**8228-8235.

43. **Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. Horner-Devine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Øvreås, A.-L. Reysenbach, V. H. Smith, and J. T. Staley.** 2006. Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. **4:**102-112.

44. **McBeth, J. M., B. J. Little, R. I. Ray, K. M. Farrar, and D. Emerson.** 2011. Neutrophilic iron-oxidizing "*Zetaproteobacteria*" and mild steel corrosion in nearshore marine environments. Appl. Environ. Microbiol. **77:**1405-1412.

45. **Moreau, J. W., R. A. Zierenberg, and J. F. Banfield.** 2010. Diversity of dissimilatory sulfite reductase genes (*dsrAB*) in a salt marsh impacted by long-term acid mine drainage. Appl. Environ. Microbiol. **76:**4819-4828.

46. **Moyer, C. L.** 2001. Molecular phylogeny: applications and implications for marine microbiology. Meth. Microbiol. **30:**375-394.

47. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1994. Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **60:**871-879.

48. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1995. Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **61:**1555-1562.

49. **Oden, N. L., and R. R. Sokal.** 1986. Directional autocorrelation: an extension of spatial correlograms to two dimensions. Syst. Zool. **35:**608-617.

50. **Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek.** 1994. fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. **10:**41-48.

51. **O'Malley, M. A.** 2007. The nineteenth century roots of 'everything is everywhere.' Nat. Rev. Microbiol. **5:**647-651.

52. **Papke, R. T., N. B. Ramsing, M. M. Bateson, and D. M. Ward.** 2003. Geographical isolation in hot spring cyanobacteria. Environ. Microbiol. **5:**650-659.

53. **Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35:**7188-7196.

54. **Ramette, A., and J. M. Tiedje.** 2007. Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. Microb. Ecol. **53:**197-207.

55. **Rassa, A. C., S. M. McAllister, S. A. Safran, and C. L. Moyer.** 2009. *Zeta-Proteobacteria* dominate the colonization and formation of microbial mats in low-temperature hydrothermal vents at Loihi Seamount, Hawaii. Geomicrobiol. J. **26:**623-638.

56. **Santelli, C. M., B. N. Orcutt, E. Banning, W. Bach, C. L. Moyer, M. L. Sogin, H. Staudigel, and K. J. Edwards.** 2008. Abundance and diversity of microbial life in ocean crust. Nature **453:**653-656.

57. **Schloss, P. D.** 2008. Evaluating different approaches that test whether microbial communities have the same structure. ISME J. **2:**265-275.

58. **Schloss, P. D., and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. **71:**1501-1506.

59. **Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. **75:**7537-7541.

60. **Shank, T. M.** 2010. Seamounts: deep-ocean laboratories of faunal connectivity, evolution, and endemism. Oceanography **23:**108-122.

61. **Sobolev, D., and E. E. Roden.** 2004. Characterization of a neutrophilic, chemolithoautotrophic Fe(II)-oxidizing β-Proteobacterium from freshwater wetland sediments. Geomicrobiol. J. **21:**1-10.

62. **Sokal, R. R.** 1986. Spatial data analysis and historical processes, p. 29-43. *In* E. Diday *et al.* (ed.), Data analysis and informatics, IV. Elsevier Science Pub. Co., Amsterdam, North-Holland.

63. **Southam, G.** 2000. Bacterial surface-mediated mineral formation, p. 257-276. *In* D. R. Lovley (ed.), Environmental microbe-metal interactions. ASM Press, Washington, D.C.

64. **Sudek, L. A., A. S. Templeton, B. M. Tebo, and H. Staudigel.** 2009. Microbial ecology of Fe (hydr)oxide mats and basaltic rock from Vailulu'u Seamount, American Samoa. Geomicrobiol. J. **26:**581-596.

65. **Takai, K., S. Nakagawa, A.-L. Reysenbach, and J. Hoek.** 2006. Microbial ecology of mid-ocean ridges and back-arc basins, p. 185-213. *In* D. M. Christie, C. R. Fisher, S.-M. Lee, and S. Givens (ed.), Back-arc spreading systems: geological, biological, chemical, and physical interactions. Geophysical Monograph Series 166, American Geophysical Union, Washington, D.C.

66. **Van der Gucht, K., K. Cottenie, K. Muylaert, N. Vloemans, S. Cousin, S. Declerck, E. Jeppesen, J.-M. Conde-Porcuna, K. Schwenk, G. Zwart, H. Degans, W. Vyverman, and L. De Meester.** 2007. The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. Proc. Natl. Acad. Sci. U.S.A. **104:**20404-20409.

67. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173:**697-703.

68. **Wessel, P., D. T. Sandwell, and S.-S. Kim.** 2010. The global seamount census. Oceanography **23:**24-33.

69. **Wheat, C. G., H. W. Jannasch, J. N. Plant, C. L. Moyer, F. J. Sansone, and G. M. McMurtry.** 2000. Continuous sampling of hydrothermal fluids from Loihi Seamount after the 1996 event. J. Geophys. Res. **105:**19353-19367.

70. **Whitaker, R. J., D. W. Grogan, and J. W. Taylor.** 2003. Geographic barriers isolate endemic populations of hyperthermophilic Archaea. Science **301:**976-978.

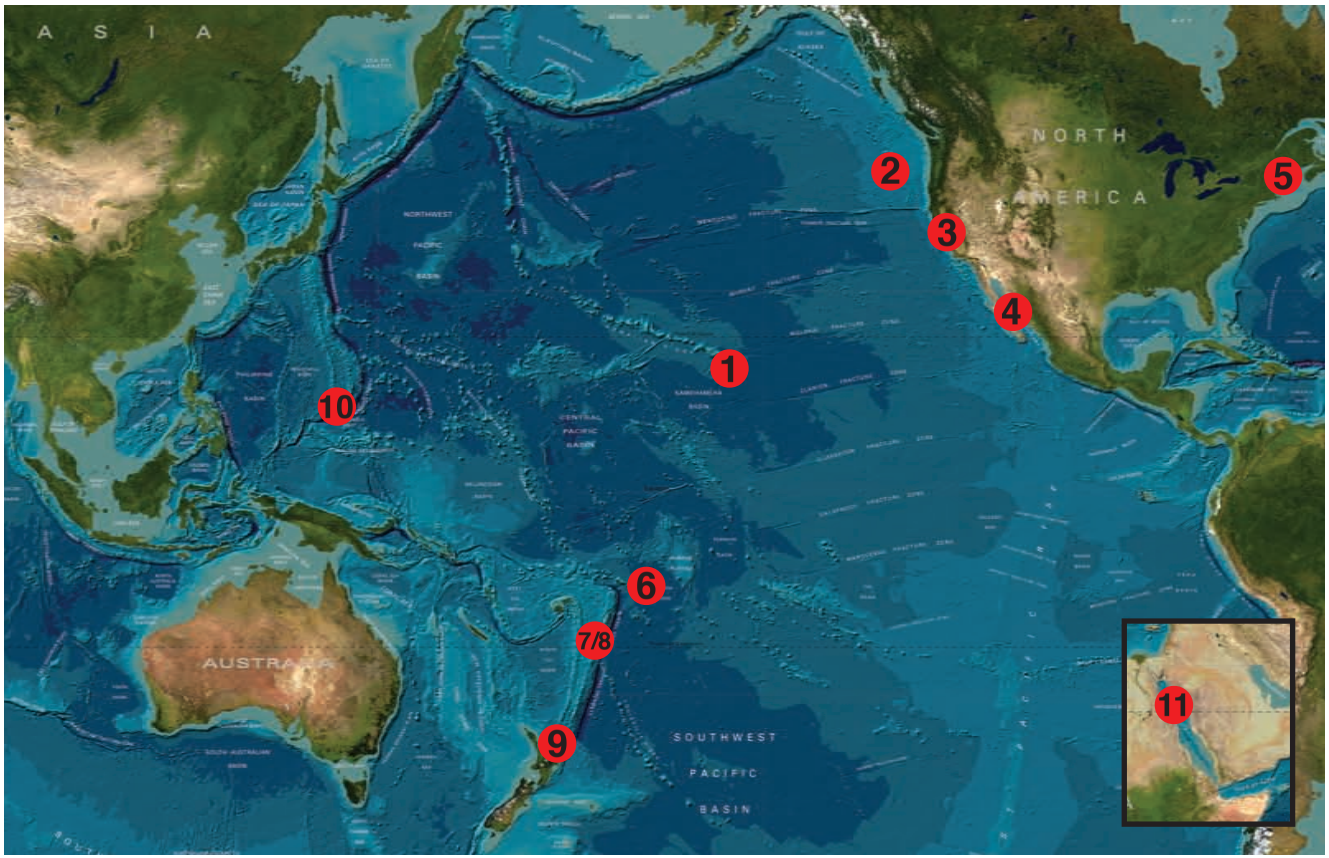


FIG. 1. Bathymetric map showing the eleven regions where full-length *Zetaproteobacteria* sequences have been detected. These include: 1) Loihi Seamount, 2) Juan de Fuca Ridge, 3) San Francisco Bay, California, 4) Southern Guaymas vent field, 5) coastal Maine, 6) Vailulu'u Seamount, 7) East Lau Spreading Center, 8) Tonga Arc, 9) Kermadec Arc, 10) Southern Mariana Trough, and 11) Red Sea. Image reproduced with permission from the GEBCO world map, <http://www.gebco.net/>.

FIG. 2. Maximum likelihood phylogenetic tree showing the evolutionary placement of all 84 full-length *Zetaproteobacteria* sequences used in this study (1282 bp mask). Red, blue, and green color groupings indicate clones from the central (Loihi Seamount), southern (Vailulu'u Seamount/Tonga Arc/ELSC/Kermadec Arc), and western (Southern Mariana Trough) Pacific Ocean, respectively. Novel sequences from this study are highlighted. The top eleven OTUs are indicated, along with the borehole fluid endemic, OTU 15. Accession numbers for published sequences are shown in parentheses in addition to the number of clones represented by each sequence. Only bootstrap values above 50 are shown. Scale bar represents 5 nucleotide substitutions per 100 positions.

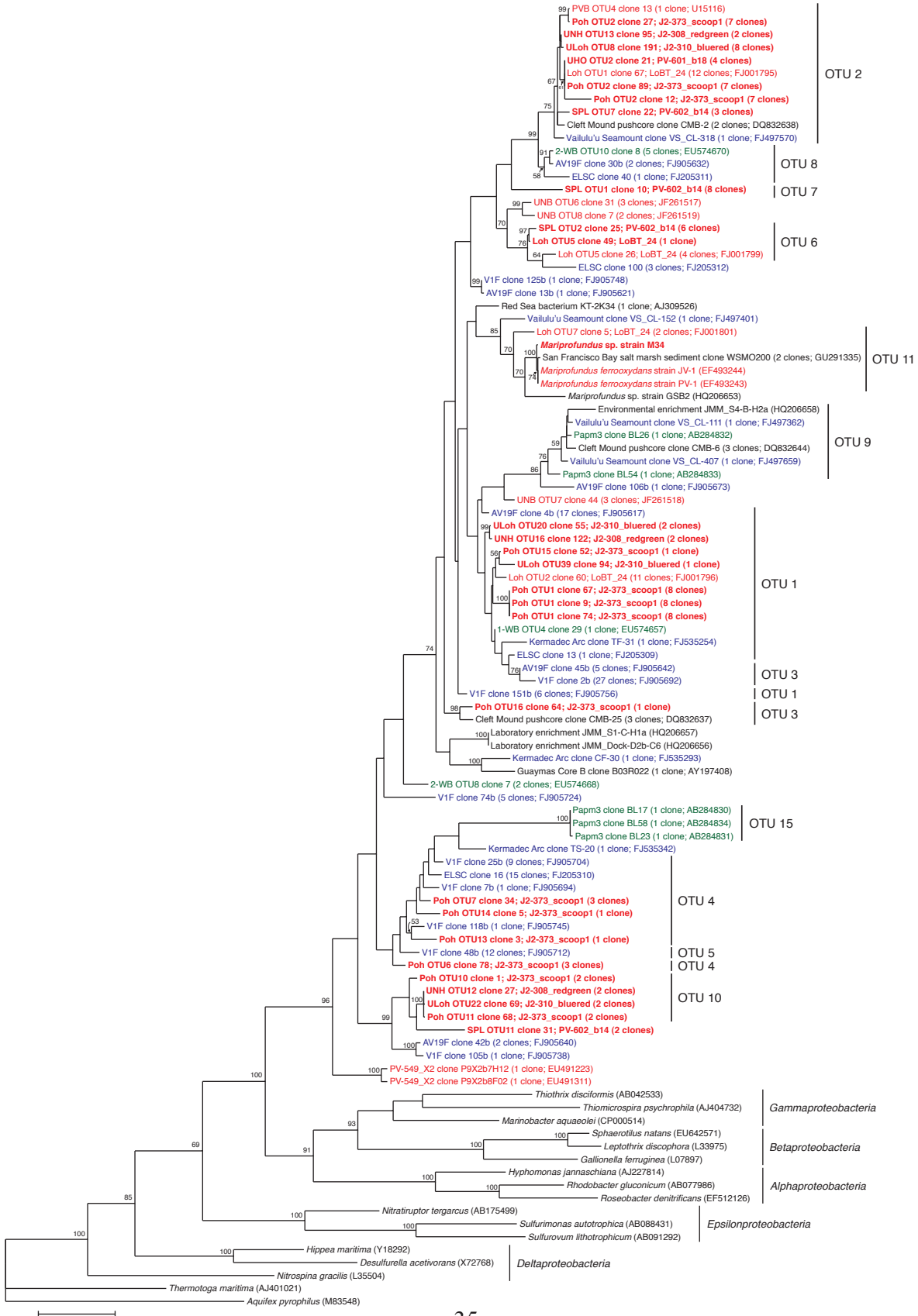

PVB OTU4 clone 13 (1 clone; U15116)
Poh OTU2 clone 27; J2-373_scoop1 (7 clones)
UNH OTU13 clone 95; J2-308_redgreen (2 clones)
ULoh OTU8 clone 191; J2-310_bluered (8 clones)
UHO OTU2 clone 21; PV-601_b18 (4 clones)
Loh OTU1 clone 67; LoBT_24 (12 clones; FJ001795)
Poh OTU2 clone 89; J2-373_scoop1 (7 clones)
Poh OTU2 clone 12; J2-373_scoop1 (7 clones)
SPL OTU7 clone 22; PV-602_b14 (3 clones)
Cleft Mound pushcore clone CMB-2 (2 clones; DQ832638)
Vailulu'u Seamount clone VS_CL-318 (1 clone; FJ497570)
OTU 2
2-WB OTU10 clone 8 (5 clones; EU574670)
AV19F clone 30b (2 clones; FJ905632)
ELSC clone 40 (1 clone; FJ205311)
OTU 8
SPL OTU1 clone 10; PV-602_b14 (8 clones)
OTU 7
UNB OTU6 clone 31 (3 clones; JF261517)
UNB OTU8 clone 7 (2 clones; JF261519)
SPL OTU2 clone 25; PV-602_b14 (6 clones)
Loh OTU5 clone 49; LoBT_24 (1 clone)
Loh OTU5 clone 26; LoBT_24 (4 clones; FJ001799)
ELSC clone 100 (3 clones; FJ205312)
OTU 6
V1F clone 125b (1 clone; FJ905748)
AV19F clone 13b (1 clone; FJ905621)
Red Sea bacterium KT-2K34 (1 clone; AJ309526)
Vailulu'u Seamount clone VS_CL-152 (1 clone; FJ497401)
Loh OTU7 clone 5; LoBT_24 (2 clones; FJ001801)
Mariprofundus sp. strain M34
San Francisco Bay salt marsh sediment clone WSMO200 (2 clones; GU291335)
Mariprofundus ferrooxydans strain JV-1 (EF493244)
Mariprofundus ferrooxydans strain PV-1 (EF493243)
OTU 11
Mariprofundus sp. strain GSB2 (HQ206653)
Environmental enrichment JMM_S4-B-H2a (HQ206658)
Vailulu'u Seamount clone VS_CL-111 (1 clone; FJ497362)
Papm3 clone BL26 (1 clone; AB284832)
Cleft Mound pushcore clone CMB-6 (3 clones; DQ832644)
Vailulu'u Seamount clone VS_CL-407 (1 clone; FJ497659)
Papm3 clone BL54 (1 clone; AB284833)
AV19F clone 106b (1 clone; FJ905673)
OTU 9
UNB OTU7 clone 44 (3 clones; JF261518)
AV19F clone 4b (17 clones; FJ905617)
ULoh OTU20 clone 55; J2-310_bluered (2 clones)
UNH OTU16 clone 122; J2-308_redgreen (2 clones)
Poh OTU15 clone 52; J2-373_scoop1 (1 clone)
ULoh OTU39 clone 94; J2-310_bluered (1 clone)
Loh OTU2 clone 60; LoBT_24 (11 clones; FJ001796)
Poh OTU1 clone 67; J2-373_scoop1 (8 clones)
Poh OTU1 clone 9; J2-373_scoop1 (8 clones)
Poh OTU1 clone 74; J2-373_scoop1 (8 clones)
1-WB OTU4 clone 29 (1 clone; EU574657)
Kermadec Arc clone TF-31 (1 clone; FJ535254)
ELSC clone 13 (1 clone; FJ205309)
OTU 1
AV19F clone 45b (5 clones; FJ905642)
V1F clone 2b (27 clones; FJ905692)
OTU 3
V1F clone 151b (6 clones; FJ905756)
OTU 1
Poh OTU16 clone 64; J2-373_scoop1 (1 clone)
Cleft Mound pushcore clone CMB-25 (3 clones; DQ832637)
OTU 3
Laboratory enrichment JMM_S1-C-H1a (HQ206657)
Laboratory enrichment JMM_Dock-D2b-C6 (HQ206656)
Kermadec Arc clone CF-30 (1 clone; FJ535293)
Guaymas Core B clone B03R022 (1 clone; AY197408)
2-WB OTU8 clone 7 (2 clones; EU574668)
V1F clone 74b (5 clones; FJ905724)
Papm3 clone BL17 (1 clone; AB284830)
Papm3 clone BL58 (1 clone; AB284834)
Papm3 clone BL23 (1 clone; AB284831)
OTU 15
Kermadec Arc clone TS-20 (1 clone; FJ535342)
V1F clone 25b (9 clones; FJ905704)
ELSC clone 16 (15 clones; FJ205310)
V1F clone 7b (1 clone; FJ905694)
Poh OTU7 clone 34; J2-373_scoop1 (3 clones)
Poh OTU14 clone 5; J2-373_scoop1 (1 clone)
V1F clone 118b (1 clone; FJ905745)
Poh OTU13 clone 3; J2-373_scoop1 (1 clone)
OTU 4
V1F clone 48b (12 clones; FJ905712)
OTU 5
Poh OTU6 clone 78; J2-373_scoop1 (3 clones)
OTU 4
Poh OTU10 clone 1; J2-373_scoop1 (2 clones)
UNH OTU12 clone 27; J2-308_redgreen (2 clones)
ULoh OTU22 clone 69; J2-310_bluered (2 clones)
Poh OTU11 clone 68; J2-373_scoop1 (2 clones)
SPL OTU11 clone 31; PV-602_b14 (2 clones)
OTU 10
AV19F clone 42b (2 clones; FJ905640)
V1F clone 105b (1 clone; FJ905738)
PV-549_X2 clone P9X2b7H12 (1 clone; EU491223)
PV-549_X2 clone P9X2b8F02 (1 clone; EU491311)
Thiothrix disciformis (AB042533)
Thiomicrospira psychrophila (AJ404732)
Marinobacter aquaeolei (CP000514)
Gammaproteobacteria
Sphaerotilus natans (EU642571)
Leptothrix discophora (L33975)
Gallionella ferruginea (L07897)
Betaproteobacteria
Hyphomonas jannaschiana (AJ227814)
Rhodobacter gluconicum (AB077986)
Roseobacter denitrificans (EF512126)
Alphaproteobacteria
Nitratiruptor tergarcus (AB175499)
Sulfurimonas autotrophica (AB088431)
Sulfurovum lithotrophicum (AB091292)
Epsilonproteobacteria
Hippea maritima (Y18292)
Desulfurella acetivorans (X72768)
Deltaproteobacteria
Nitrospina gracilis (L35504)
Thermotoga maritima (AJ401021)
Aquifex pyrophilus (M83548)
0.05

FIG. 3. Maximum likelihood phylogenetic tree showing the evolutionary placement of all 132 full- and partial-length *Zetaproteobacteria* sequences used in this study (696 bp mask). Red and green color groupings indicate clones from microbial mats and borehole fluids, respectively. Novel sequences from this study are highlighted. Selected OTUs are indicated for reference. Accession numbers for published sequences are shown in parentheses in addition to the number of clones represented by each sequence. Only bootstrap values above 50 are shown. Scale bar represents 5 nucleotide substitutions per 100 positions.


PVB OTU4 clone 13 (1 clone; U15116)
Poh OTU2 clone 27; J2-373_scoop1 (7 clones)
Loh OTU1 clone 67; LoBT_24 (12 clones; FJ001795)
Poh OTU2 clone 89; J2-373_scoop1 (7 clones)
UHO OTU2 clone 21; PV-601_b18 (4 clones)
UNH OTU13 clone 95; J2-308_redgreen (2 clones)
ULoh OTU8 clone 191; J2-310_bluered (8 clones)
ELSC clone 40 (1 clone; FJ205311)
OTU 2
2-WB OTU10 clone 8 (5 clones; EU574670)
AV19F clone 30b (2 clones; FJ905632)
Fapm1b clone B05 (5 clones; AB213124)
Fapm1b clone B31 (1 clone; AB213131)
YS18U clone s85 (2 clones; AB329981)
Vailulu'u Seamount clone VS_CL-318 (1 clone; FJ497570)
SPL OTU11 clone 31; PV-602_b14 (2 clones)
YS18U clone s36 (3 clones; AB329965)
YS18U clone c23 (1 clone; AB329934)
Cleft Mound pushcore clone CMB-2 (2 clones; DQ832638)
SPL OTU7 clone 22; PV-602_b14 (3 clones)
UNB OTU6 clone 31 (3 clones; JF261517)
UNB OTU8 clone 7 (2 clones; JF261519)
Loh OTU5 clone 49; LoBT_24 (1 clone)
SPL OTU2 clone 25; PV-602_b14 (6 clones)
YS18U clone c21 (1 clone; AB329933)
YS18U clone s21 (3 clones; AB329959)
ELSC clone 100 (3 clones; FJ205312)
SPL OTU1 clone 10; PV-602_b14 (8 clones)
YS18U clone c19 (1 clone; AB329931)
YS18U clone s10 (2 clones; AB329956)
V1F clone 125b (1 clone; FJ905748)
AV19F clone 13b (1 clone; FJ905621)
V1F clone 151b (6 clones; FJ905756)
YS18U clone c02 (7 clones; AB329923)
Red Sea bacterium KT-2K34 (1 clone; AJ309526)
Vailulu'u Seamount clone VS_CL-152 (1 clone; FJ497401)
San Francisco Bay salt marsh sediment clone WSMO200 (2 clones; GU291335)
Mariprofundus ferrooxydans strain PV-1 (EF493243)
Mariprofundus ferrooxydans strain JV-1 (EF493244)
Mariprofundus sp. strain M34
Loh OTU5 clone 26; LoBT_24 (4 clones; FJ001799)
YS16U clone c23 (4 clones; AB329857)
Loh OTU2 clone 60; LoBT_24 (11 clones; FJ001796)
AV19F clone 4b (17 clones; FJ905617)
YS18B clone c44 (2 clones; AB330038)
YS18U clone s02 (23 clones; AB329951)
ULoh OTU39 clone 94; J2-310_bluered (1 clone)
Poh OTU15 clone 52; J2-373_scoop1 (1 clone)
ELSC clone 13 (1 clone; FJ205309)
OTU 1
Kermadec Arc clone TF-31 (1 clone; FJ535254)
1-WB OTU4 clone 29 (1 clone; EU574657)
Fapm1a clone B29 (1 clone; AB213117)
ULoh OTU20 clone 55; J2-310_bluered (2 clones)
UNH OTU16 clone 122; J2-308_redgreen (2 clones)
V1F clone 2b (27 clones; FJ905692)
M48 clone B02; J2-373_scoop5 (1 clone; JF440632)
YS18B clone c63 (2 clones; AB330049)
Poh OTU2 clone 12; J2-373_scoop1 (7 clones)
Poh OTU1 clone 74; J2-373_scoop1 (8 clones)
Poh OTU1 clone 9; J2-373_scoop1 (8 clones)
Poh OTU1 clone 67; J2-373_scoop1 (8 clones)
OTU 1
M48 clone D03; J2-373_scoop5 (1 clone; JF440633)
M48 clone A04; J2-373_scoop5 (1 clone; JF440627)
AV19F clone 45b (5 clones; FJ905642)
Loh OTU7 clone 5; LoBT_24 (2 clones; FJ001801)
Mariprofundus sp. strain GSB2 (HQ206653)
UNB OTU7 clone 44 (3 clones; JF261518)
YS16U clone c25 (1 clone; AB329859)
Poh OTU16 clone 64; J2-373_scoop1 (1 clone)
Cleft Mound pushcore clone CMB-25 (3 clones; DQ832637)
Laboratory enrichment JMM_S1-C-H1a (HQ206657)
Laboratory enrichment JMM_Dock-D2b-C6 (HQ206656)
YS18U clone s07 (2 clones; AB329954)
M48 clone D01; J2-373_scoop5 (1 clone; JF440630)
V1F clone 25b (9 clones; FJ905704)
Poh OTU13 clone 3; J2-373_scoop1 (1 clone)
ELSC clone 16 (15 clones; FJ205310)
M48 clone B04; J2-373_scoop5 (1 clone; JF440629)
OTU 4
M48 clone A11; J2-373_scoop5 (1 clone; JF440628)
V1F clone 7b (1 clone; FJ905694)
V1F clone 118b (1 clone; FJ905745)
Poh OTU14 clone 5; J2-373_scoop1 (1 clone)
Poh OTU7 clone 34; J2-373_scoop1 (3 clones)
V1F clone 48b (12 clones; FJ905712)
M48 clone F12; J2-373_scoop5 (1 clone; JF440635)
Poh OTU6 clone 78; J2-373_scoop1 (3 clones)
V1F clone 105b (1 clone; FJ905738)
AV19F clone 42b (2 clones; FJ905640)
ULoh OTU22 clone 69; J2-310_bluered (2 clones)
M48 clone E06; J2-373_scoop5 (1 clone; JF440631)
Poh OTU11 clone 68; J2-373_scoop1 (2 clones)
UNH OTU12 clone 27; J2-308_redgreen (2 clones)
Poh OTU10 clone 1; J2-373_scoop1 (2 clones)
M48 clone D06; J2-373_scoop5 (1 clone; JF440634)
YS18U clone c08 (1 clone; AB329927)
YS16U clone s17 (1 clone; AB329888)
YS16U clone c03 (3 clones; AB329845)
PV-549_X2 clone P9X2b8F02 (1 clone; EU491311)
PV-549_X2 clone P9X2b7H12 (1 clone; EU491223)
2-WB OTU8 clone 7 (2 clones; EU574668)
V1F clone 74b (5 clones; FJ905724)
YS18U clone c33 (1 clone; AB329938)
Guaymas Core B clone B03R022 (1 clone; AY197408)
YS16U clone c02 (1 clone; AB329844)
Kermadec Arc clone CF-30 (1 clone; FJ535293)
YS18U clone s09 (2 clones; AB329955)
YS18U clone c24 (4 clones; AB329935)
Kermadec Arc clone TS-20 (1 clone; FJ535342)
Fapm1b clone B60 (1 clone; AB213138)
OTU 9
Fapm1a clone B23 (1 clone; AB213115)
AV19F clone 106b (1 clone; FJ905673)
Environmental enrichment JMM_S4-B-H2a (HQ206658)
Fapm1b clone B25 (1 clone; AB213129)
Fapm1b clone B09 (3 clones; AB213125)
Papm3 clone B23 (1 clone; AB213181)
Fapm1a clone B03 (4 clones; AB213109)
Vailulu'u Seamount clone VS_CL-407 (1 clone; FJ497659)
Cleft Mound pushcore clone CMB-6 (3 clones; DQ832644)
OTU 9
Papm3 clone B26 (1 clone; AB213183)
Papm3 clone BL54 (1 clone; AB284833)
Fapm1a clone B33 (1 clone; AB213119)
Papm3 clone B09 (1 clone; AB213177)
Papm3 clone B02 (13 clones; AB213174)
Vailulu'u Seamount clone VS_CL-111 (1 clone; FJ497362)
Papm3 clone B06 (1 clone; AB213176)
Papm3 clone BL26 (1 clone; AB284832)
Papm3 clone BL58 (1 clone; AB284834)
Papm3 clone BL17 (1 clone; AB284830)
Papm3 clone B25 (1 clone; AB213182)
Papm3 clone BL23 (1 clone; AB284831)
OTU 15
Papm3 clone B11 (11 clones; AB213178)
Papm3 clone B31 (1 clone; AB213184)
YS16U clone c27 (1 clone; AB329861)
Marinobacter aquaeolei (CP000514)
Thiomicrospira psychrophila (AJ404732)
Gammaproteobacteria
Thiothrix disciformis (AB042533)
Sphaerotilus natans (EU642571)
Leptothrix discophora (L33975)
Betaproteobacteria
Gallionella ferruginea (L07897)
Hyphomonas jannaschiana (AJ227814)
Rhodobacter gluconicum (AB077986)
Alphaproteobacteria
Roseobacter denitrificans (EF512126)
Nitratiruptor tergarcus (AB175499)
Sulfurimonas autotrophica (AB088431)
Epsilonproteobacteria
Sulfurovum lithotrophicum (AB091292)
Hippea maritima (Y18292)
Desulfurella acetivorans (X72768)
Deltaproteobacteria
Nitrospina gracilis (L35504)
Thermotoga maritima (AJ401021)
Aquifex pyrophilus (M83548)
0.05

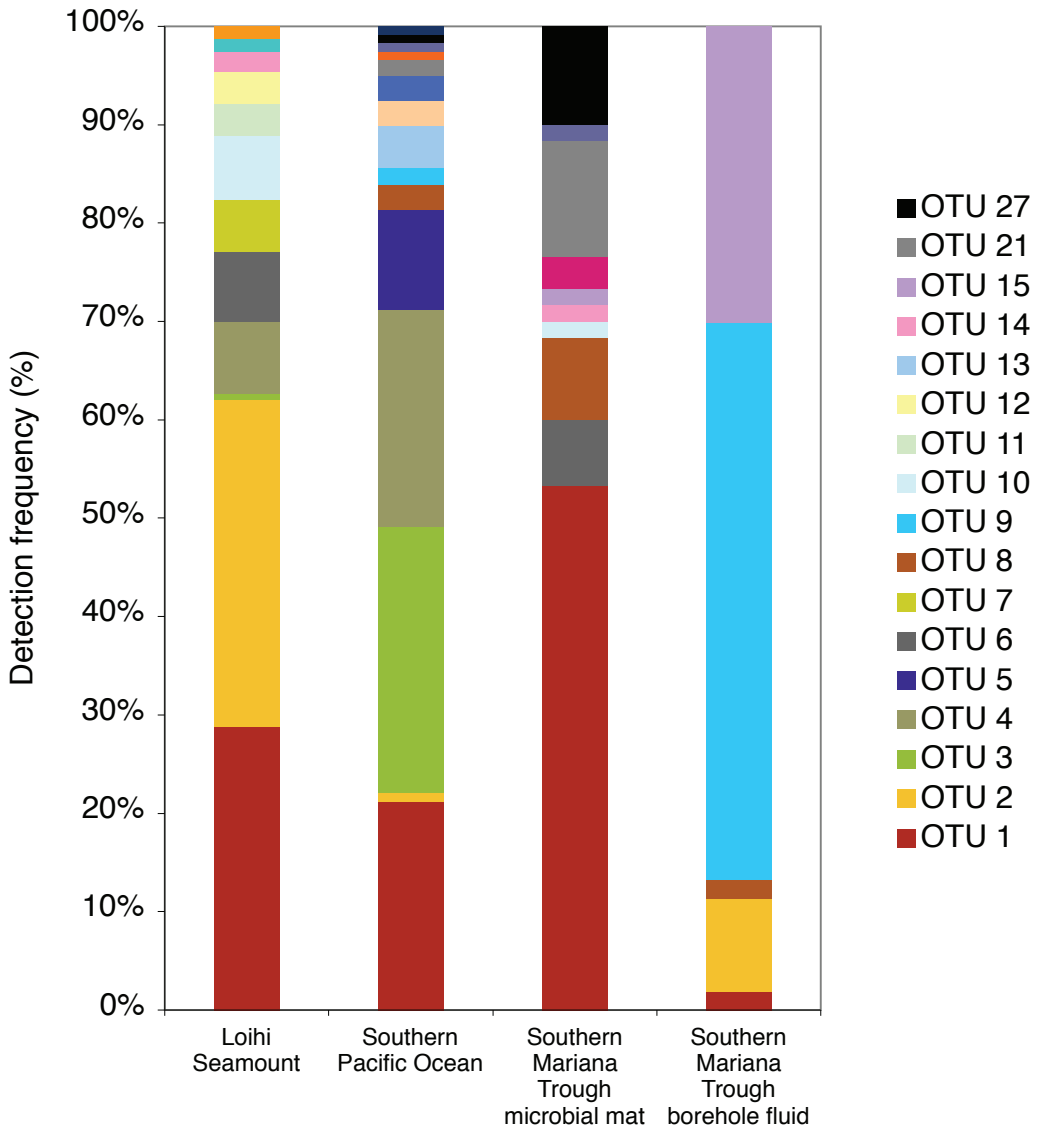


FIG. 4. Stacked bar graph showing OTU distribution within the three main sampling regions, with the Southern Mariana Trough separated by sample type. Full- plus partial-length sequence dataset used.

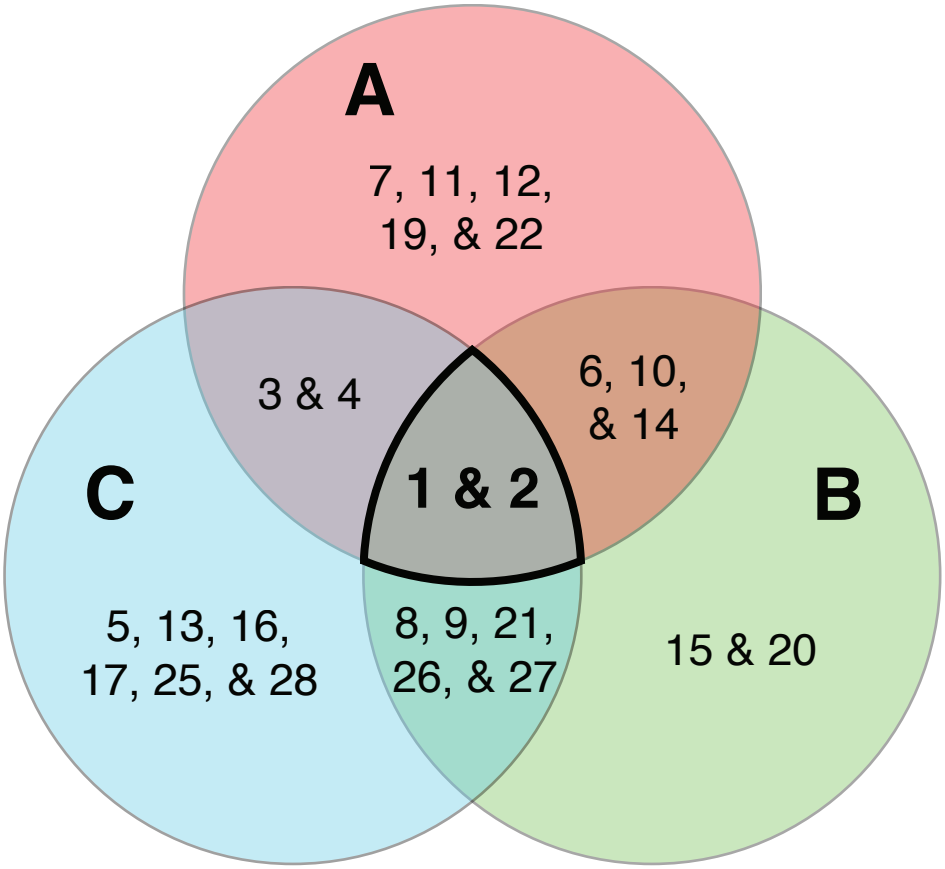


FIG. 5. Venn diagram comparing OTU distribution between A) Loihi Seamount, B) the Southern Mariana Trough, and C) the southern Pacific Ocean group (Vailulu'u Seamount/Tonga Arc/ELSC/Kermadec Arc). Ubiquitous OTUs are highlighted. Full- plus partial-length sequence dataset used.

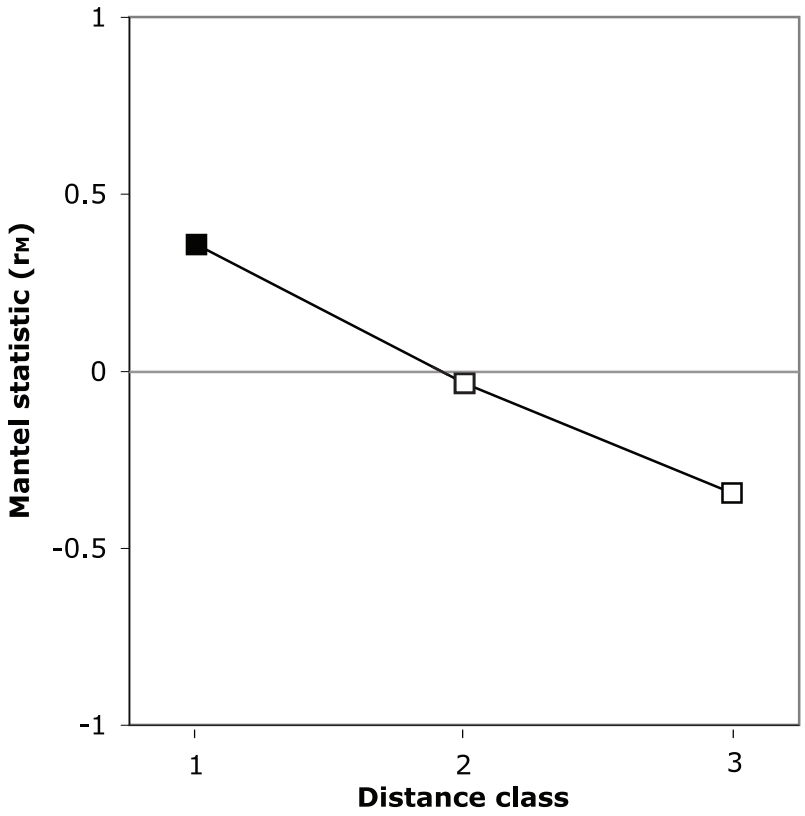


FIG. 6. Mantel correlogram showing spatial autocorrelation analysis at three distance classes using the multivariate Mantel test statistic (r_M). Significant spatial autocorrelation ($\alpha = 0.05$) indicated by a closed square, demonstrating correlation between genetic distance and spatial distance.

TABLE 1. Clone library, cultured isolate, sample type, and temperature data for *Zetaproteobacteria* SSU rRNA gene sequences used in this study.

Year Collected	Clone Library/ Cultured Isolate[1]	Region	Site	#ζ Clones[2] per Library	%ζ Clones[2] per Library	Sample Type	Temp. (˚C)	Reference
Full-Length Sequences								
1991	Pele's Vents Bacteria (PVB)	Loihi Seamount	Pele's Vents	1	<5	microbial mat	37	Moyer *et al.*, 1994; Moyer *et al.*, 1995
1996	*M. ferrooxydans* strain PV-1	Loihi Seamount	Naha Vents (Mkr #3-6)	n.a.	n.a.	microbial mat	23	Emerson and Moyer, 2002; Emerson *et al.*, 2007
1998	*M. ferrooxydans* strain JV-1	Loihi Seamount	Ikaika (Mkr #11)	n.a.	n.a.	microbial mat	165	Emerson and Moyer, 2002; Emerson *et al.*, 2007
2003	PV-549_X2	Loihi Seamount	Pisces Peak	2	<5	altered basalt	4	Santelli *et al.*, 2008
2004	**PV-601_b18 (UHO)**	Loihi Seamount	Upper Hiolo (Mkr #36)	4	8	microbial mat	57	**This study**
2004	**PV-602_b14 (SPL)**	Loihi Seamount	Spillway (Mkr #34)	19	38	microbial mat	63	**This study**
2006	***Mariprofundus* sp. strain M34**	Loihi Seamount	Spillway (Mkr #34)	n.a.	n.a.	microbial mat	52	**This study**
2006	Growth Chamber LoBT_24 (Loh)	Loihi Seamount	Lohiau (Mkr #2-5)	30	62	growth chamber	22	Rassa *et al.*, 2009
2006	Ula Nui Bacteria (UNB)	Loihi Seamount	Ula Nui (FeMO Deep Site)	8	10	manganese crust	1.7	Edwards *et al.*, 2011
2007	**J2-308_redgreen (UNH)**	Loihi Seamount	Upper N Hiolo (Mkr #39)	6	5.3	microbial mat	53	**This study**
2007	**J2-310_bluered (ULoh)**	Loihi Seamount	Upper Lohiau (Mkr #55)	13	8.8	microbial mat	22	**This study**
2008	**J2-373_scoop1 (Poh)**	Loihi Seamount	Pohaku (Mkr #57)	59	70	microbial mat	27	**This study**
1997	Red Sea KT-2	Red Sea	Kebrit Deep	1	7	brine/seawater interface	22	Eder *et al.*, 2001
1998	Guaymas Core B	Southern Guaymas vent field	Everest Mound	1	<5	sediment core	3-16	Dhillon *et al.*, 2003
2002	Cleft Mound pushcore 23	Juan de Fuca Ridge	Cleft Segment (off-axis)	8	15	sediment core	5.6	Davis *et al.*, 2009
2003	WSMO200	California	San Francisco Bay (salt marsh)	2	<5	surface sediment	10-18	Moreau *et al.*, 2010
2003	1-WB	Southern Mariana Trough	Fryer Site	1	<5	microbial mat	77	Davis and Moyer, 2008
2003	2-WB	Southern Mariana Trough	Fryer Site	7	12	microbial mat	77	Davis and Moyer, 2008
2004	Papm3	Southern Mariana Trough	Pika Site	5	50	borehole fluid	6-12	Kato *et al.*, 2009b
2005	Tangaroa Floc (TF)	Kermadec Arc	Tangaroa Seamount	1	<5	microbial mat	14.2	Hodges and Olson, 2009
2005	Tangaroa Sediment (TS)	Kermadec Arc	Tangaroa Seamount	1	<5	deep-sea sediment	12	Hodges and Olson, 2009
2005	Clark Floc (CF)	Kermadec Arc	Clark Seamount	1	<5	microbial mat	6.1	Hodges and Olson, 2009
2005	Vailulu'u Seamount (VS_CL)	Vailulu'u Seamount	Nafanua summit	4	<5	microbial mat	5.8	Sudek *et al.*, 2009
2007	East Lau Spreading Center (ELSC)	ELSC	TVG9 (near Tui Malila)	20	16	sediment core	>41.2	GenBank FJ205309-FJ205312; C. Dong, pers. comm.
2007	R1053 (V1F)	Tonga Arc	Volcano 1	63	43	microbial mat	17.2	Forget *et al.*, 2010
2007	R1046 (AV19F)	Tonga Arc	Volcano 19	28	17	microbial mat	16.2	Forget *et al.*, 2010
2008	*Mariprofundus* sp. strain GSB2	Maine	Great Salt Bay (salt marsh)	n.a.	n.a.	surface sediment	11.5	McBeth *et al.*, 2011
2009/10	Bigelow Enrichment Experiments	Maine	Boothbay Harbor/Southport Island	n.a.	n.a.	enrichment	2-21	McBeth *et al.*, 2011
Partial-Length Sequences								
2004	Papm3	Southern Mariana Trough	Pika Site	30	50	borehole fluid	6-12	Kato *et al.*, 2009b
2004	Fapm1a	Southern Mariana Trough	Fryer Site	7	11	borehole fluid	27-30	Kato *et al.*, 2009b
2004	Fapm1b	Southern Mariana Trough	Fryer Site	11	15	borehole fluid	17-27	Kato *et al.*, 2009b
2005	YS16U	Southern Mariana Trough	Kaiko Site	12	8.4	microbial mat	33	Kato *et al.*, 2009a
2005	YS18U	Southern Mariana Trough	Fryer Site	57	45	microbial mat	36-116	Kato *et al.*, 2009a
2008	J2-373_scoop5 (M48)	Loihi Seamount	Mkr #48	9	11	microbial mat	38	GenBank JF440627-JF440635

[1]Novel cultured isolate and clone libraries constructed for this study are highlighted in bold.
[2]*Zetaproteobacteria* clones
n.a. = not applicable

TABLE 2. Analysis of molecular variance (AMOVA). Significant P-values highlighted ($\alpha = 0.05$).

Sequence Subset	Grouped by:	Percentage of Variation			Among groups	
		Among groups	Among clone libraries within groups	Within clone libraries	d.f.	P-value
All Sequences	Region	8.44	25.16	66.40	7	**0.017±0.001**
	Temperature	2.67	29.89	67.44	2	0.102±0.003
	Sample Type (F)[1]	6.28	28.09	65.63	4	**0.046±0.002**
	Sample Type (F&P)[2]	15.81	21.76	62.44	4	**0.000±0.000**
Microbial Mat Samples Only	Region	8.31	19.85	71.84	2	**0.028±0.001**
	Temperature	4.71	23.35	71.94	2	0.124±0.003
OTU 1	Region	12.63	82.94	4.43	2	0.118±0.003
	Temperature	-11.98	106.87	5.11	1	0.939±0.003
OTU 2	Region	29.96	30.52	39.52	2	0.107±0.003
	Temperature	2.41	45.00	52.59	2	0.232±0.004
Subset with known [Fe]	Region	3.33	21.30	75.37	2	0.245±0.004
	Temperature	4.20	19.61	76.19	2	0.161±0.004
	total Fe (μM)	-3.47	25.54	77.93	3	0.371±0.005

[1]Full-length dataset only
[2]Full- and partial-length dataset

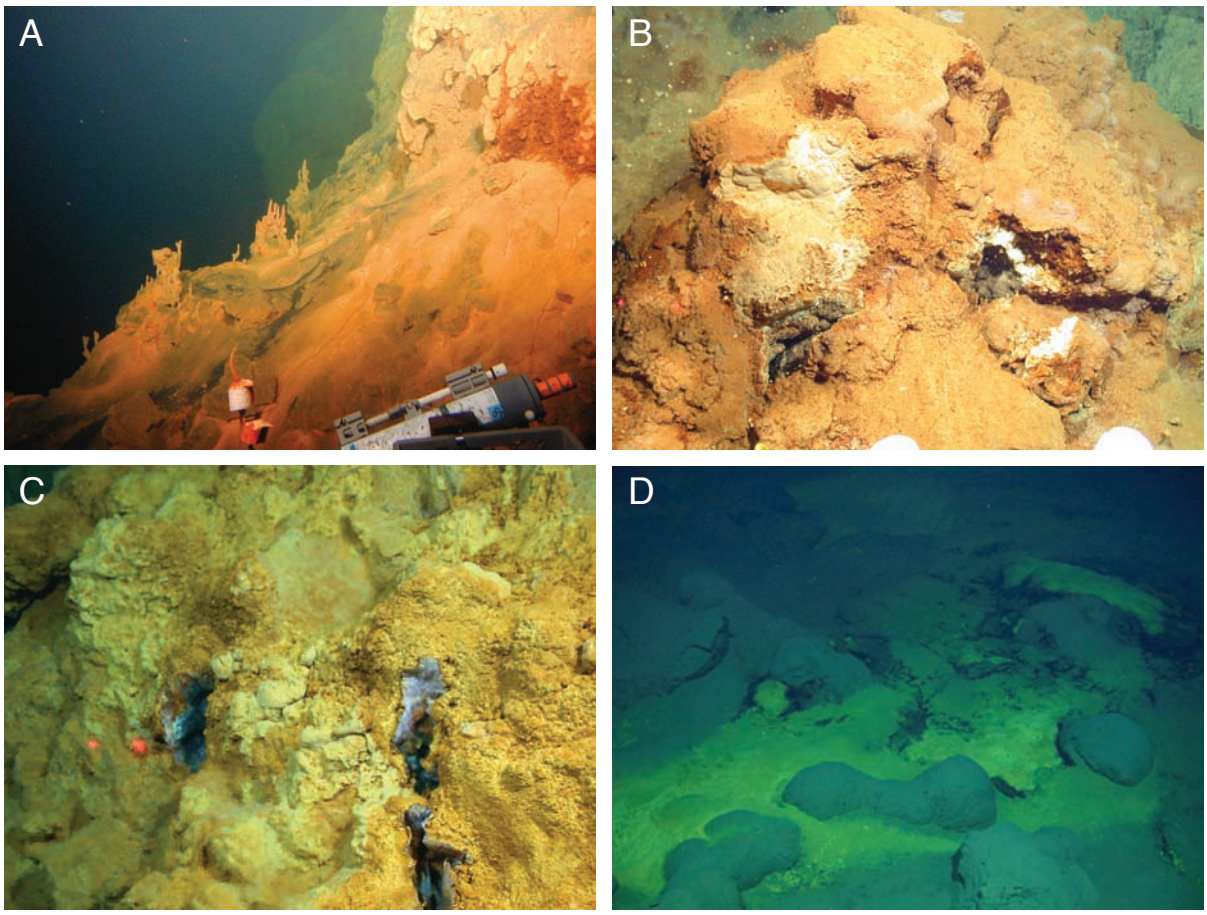


FIG. S1. Iron-oxide encrusted mats from selected sampling sites at Loihi Seamount. A) Hiolo Ridge chimlets (Upper Mkr #48). B) Pohaku (Mkr #57). C) Upper North Hiolo (Mkr #39). D) Diffuse venting at Ula Nui (FeMO Deep Site). The distance between red laser points represents 10 cm.

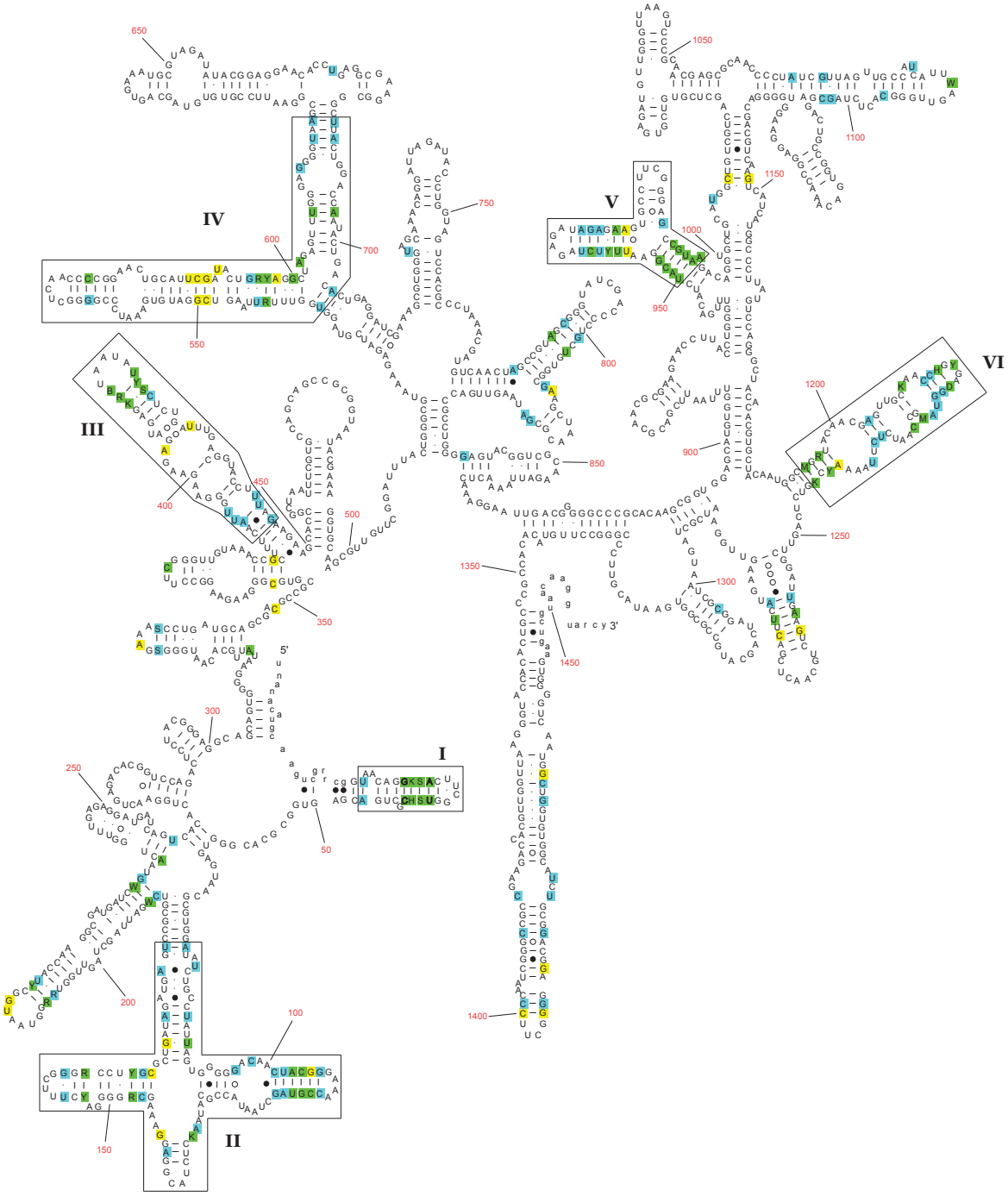


FIG. S2. SSU rRNA secondary structure analysis of the consensus sequence for *Zetaproteobacteria* OTU 1. Variability between OTU 1 and the consensus sequence for OTU 2, OTU 15, or OTUs 2 and 15, is indicated by a yellow, blue, or green highlighted base, respectively. Six regions with relatively high variability are identified.

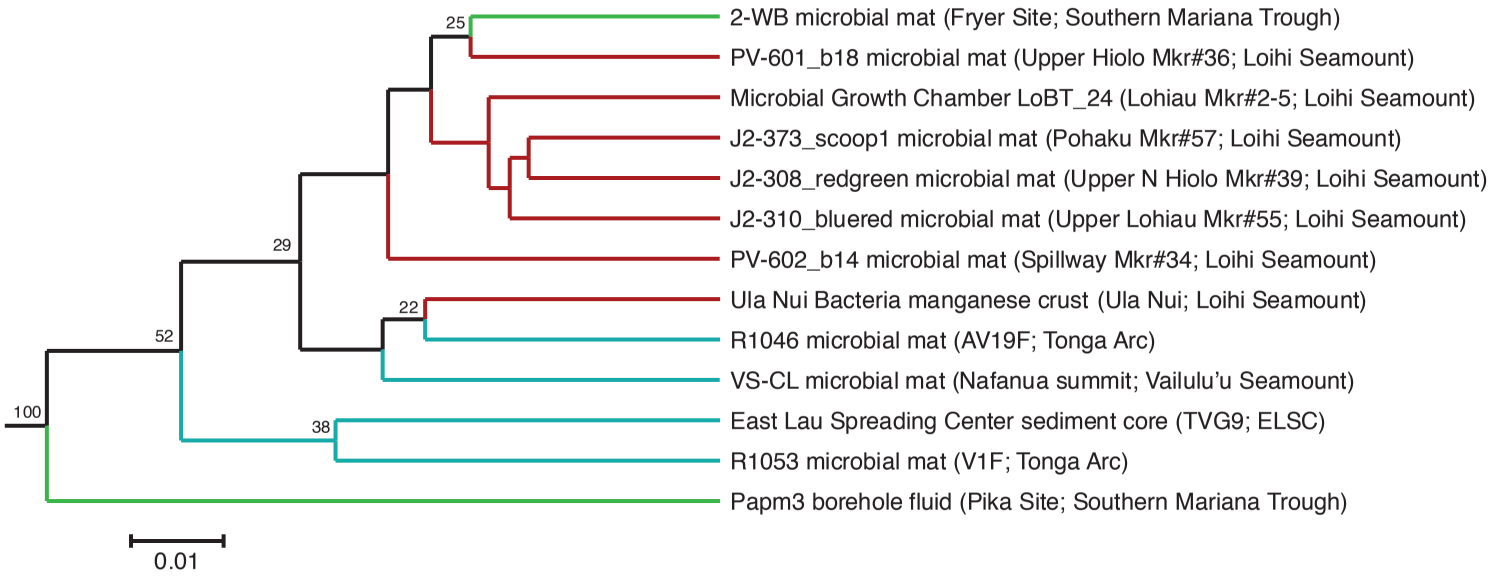


FIG. S3. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis of those samples from the three main sampling regions with four or more clones represented by full-length sequences, generated by calculating pairwise abundance-weighted non-normalized UniFrac metrics. Red, green, and blue branch colorings indicate sample origin (Loihi Seamount, Southern Mariana Trough, and the south Pacific Ocean group, respectively). Only jackknife values above 20 are shown at nodes. The scale bar represents the distance between samples in UniFrac units. The UniFrac distance metric between identical samples is 0, whereas samples with no lineage overlap have a value of 0.5 UniFrac units.

TABLE S1. Operational Taxonomic Unit (OTU) determination, representative clone number, number of clones with percent of clone library, and phylogenetic grouping for the five clone libraries constructed for this study.

Clone Library	OTU	Sequenced clone no.	No. of clones (% of library)	Phylogenetic Grouping[1]
PV-601_b18 (UHO)	1	6 & 16	18 (36.0)	Nitrospira
	2	2 & 21	4 (8.0)	ζ–*Proteobacteria*
	3	150	3 (6.0)	ε–*Proteobacteria*
	4	158	3 (6.0)	γ–*Proteobacteria*
	5	202	3 (6.0)	Chloroflexi
	6	209	2 (4.0)	ε–*Proteobacteria*
	7	65	2 (4.0)	ε–*Proteobacteria*
	8	125	2 (4.0)	ε–*Proteobacteria*
	9	34	2 (4.0)	Chloroflexi
	16	139	1 (2.0)	ε–*Proteobacteria*
PV-602_b14 (SPL)	1	10	8 (16.0)	ζ–*Proteobacteria*
	2	25	6 (12.0)	ζ–*Proteobacteria*
	3	216	4 (8.0)	Flavobacteria
	4	116	4 (8.0)	α–*Proteobacteria*
	5	40	3 (6.0)	Actinobacteria
	6	17	3 (6.0)	Chloroflexi
	7	22	3 (6.0)	ζ–*Proteobacteria*
	8	7	3 (6.0)	δ–*Proteobacteria*
	9	5	2 (4.0)	γ–*Proteobacteria*
	10	28	2 (4.0)	Nitrospira
	11	31	2 (4.0)	ζ–*Proteobacteria*
	12	69	1 (2.0)	unclassified *Proteobacteria*
J2-308_redgreen (UNH)	1	113	14 (12.4)	Nitrospira
	2	6	10 (8.8)	unclassified Nitrospira
	3	77	8 (7.1)	α–*Proteobacteria*
	4	22	6 (5.3)	unclassified *Proteobacteria*
	5	136	6 (5.3)	δ–*Proteobacteria*
	6	30	6 (5.3)	Chloroflexi
	7	34	6 (5.3)	Chloroflexi
	8	110	5 (4.4)	α–*Proteobacteria*
	9	141	5 (4.4)	Actinobacteria
	10	75	4 (3.5)	Acidobacteria
	11	114	4 (3.5)	α–*Proteobacteria*
	12	27	2 (1.8)	ζ–*Proteobacteria*
	13	95	2 (1.8)	ζ–*Proteobacteria*
	14	118	2 (1.8)	WS3
	16	122	2 (1.8)	ζ–*Proteobacteria*

(*Continued on next page*)

TABLE S1. Operational Taxonomic Unit (OTU) determination, representative clone number, number of clones with percent of clone library, and phylogenetic grouping for the five clone libraries constructed for this study. (*Continued*)

Clone Library	OTU	Sequenced clone no.	No. of clones (% of library)	Phylogenetic Grouping[1]
J2-310_bluered (ULoh)	1	110	23 (15.6)	Actinobacteria
	2	159	9 (6.1)	WS3
	3	87	9 (6.1)	unclassified *Proteobacteria*
	4	160	8 (5.4)	δ–*Proteobacteria*
	5	48	7 (4.8)	δ–*Proteobacteria*
	6	161	6 (4.1)	unclassified Nitrospira
	7	177	6 (4.1)	δ–*Proteobacteria*
	8	191	8 (5.4)	ζ–*Proteobacteria*
	9	187	5 (3.4)	TM7
	10	123	5 (3.4)	α–*Proteobacteria*
	11	68	4 (2.7)	δ–*Proteobacteria*
	12	2	3 (2.0)	Actinobacteria
	13	84	3 (2.0)	unclassified Bacteria
	20	55	2 (1.4)	ζ–*Proteobacteria*
	22	69	2 (1.4)	ζ–*Proteobacteria*
	39	94	1 (0.7)	ζ–*Proteobacteria*
J2-373_scoop1 (Poh)	1	9, 67, & 74	24 (28.6)	ζ–*Proteobacteria*
	2	12, 27, & 89	21 (25.0)	ζ–*Proteobacteria*
	3	24	7 (8.3)	Acidobacteria
	4	76	4 (4.8)	γ–*Proteobacteria*
	5	48	3 (3.6)	Acidobacteria
	6	78	3 (3.6)	ζ–*Proteobacteria*
	7	34	3 (3.6)	ζ–*Proteobacteria*
	8	10	2 (2.4)	γ–*Proteobacteria*
	9	82	2 (2.4)	γ–*Proteobacteria*
	10	1	2 (2.4)	ζ–*Proteobacteria*
	11	68	2 (2.4)	ζ–*Proteobacteria*
	13	3	1 (1.2)	ζ–*Proteobacteria*
	14	5	1 (1.2)	ζ–*Proteobacteria*
	15	52	1 (1.2)	ζ–*Proteobacteria*
	16	64	1 (1.2)	ζ–*Proteobacteria*

[1]As determined by Ribosomal Database Project (RDP) release 10.22.

TABLE S2. Operational Taxonomic Unit (OTU) designations for the full-length *Zetaproteobacteria* dataset.

OTU	Clone	Accession Number	Rep. no. clones	Region	Reference
1	Loh OTU2 clone 60	FJ001796	11	Loihi Seamount	Rassa *et al.*, 2009
1	Poh OTU1 clone 9	JF320714	8	Loihi Seamount	This study
1	Poh OTU1 clone 67	JF320715	8	Loihi Seamount	This study
1	Poh OTU1 clone 74	JF320716	8	Loihi Seamount	This study
1	Poh OTU15 clone 52	JF320731	1	Loihi Seamount	This study
1	ULoh OTU20 clone 55	JF320770	2	Loihi Seamount	This study
1	ULoh OTU39 clone 94	JF320772	1	Loihi Seamount	This study
1	UNH OTU16 clone 122	JF320787	2	Loihi Seamount	This study
1	V1F clone 151b	FJ905756	6	Tonga Arc	Forget *et al.*, 2010
1	AV19F clone 4b	FJ905617	17	Tonga Arc	Forget *et al.*, 2010
1	ELSC clone 13	FJ205309	1	East Lau Spreading Center	GenBank; C. Dong, pers. comm.
1	Kermadec Arc clone TF-31	FJ535254	1	Kermadec Arc	Hodges and Olson, 2009
1	1-WB OTU4 clone 29	EU574657	1	Southern Mariana Trough	Davis and Moyer, 2008
2	PVB OTU4 clone 13	U15116	1	Loihi Seamount	Moyer *et al.*, 1995
2	UHO OTU2 clone 21	JF320735	4	Loihi Seamount	This study
2	SPL OTU7 clone 22	JF320751	3	Loihi Seamount	This study
2	Loh OTU1 clone 67	FJ001795	12	Loihi Seamount	Rassa *et al.*, 2009
2	Poh OTU2 clone 12	JF320717	7	Loihi Seamount	This study
2	Poh OTU2 clone 27	JF320718	7	Loihi Seamount	This study
2	Poh OTU2 clone 89	JF320719	7	Loihi Seamount	This study
2	ULoh OTU8 clone 191	JF320764	8	Loihi Seamount	This study
2	UNH OTU13 clone 95	JF320785	2	Loihi Seamount	This study
2	Cleft Mound pushcore clone CMB-2	DQ832638	2	Juan de Fuca Ridge	Davis *et al.*, 2009
2	Vailulu'u Seamount clone VS_CL-318	FJ497570	1	Vailulu'u Seamount	Sudek *et al.*, 2009
3	Poh OTU16 clone 64	JF320732	1	Loihi Seamount	This study
3	Cleft Mound pushcore clone CMB-25	DQ832637	3	Juan de Fuca Ridge	Davis *et al.*, 2009
3	AV19F clone 45b	FJ905642	5	Tonga Arc	Forget *et al.*, 2010
3	V1F clone 2b	FJ905692	27	Tonga Arc	Forget *et al.*, 2010
4	Poh OTU6 clone 78	JF320723	3	Loihi Seamount	This study
4	Poh OTU7 clone 34	JF320724	3	Loihi Seamount	This study
4	Poh OTU13 clone 3	JF320729	1	Loihi Seamount	This study
4	Poh OTU14 clone 5	JF320730	1	Loihi Seamount	This study
4	V1F clone 118b	FJ905745	1	Tonga Arc	Forget *et al.*, 2010
4	V1F clone 7b	FJ905694	1	Tonga Arc	Forget *et al.*, 2010
4	V1F clone 25b	FJ905704	9	Tonga Arc	Forget *et al.*, 2010
4	ELSC clone 16	FJ205310	15	East Lau Spreading Center	GenBank; C. Dong, pers. comm.
5	V1F clone 48b	FJ905712	12	Tonga Arc	Forget *et al.*, 2010
6	SPL OTU2 clone 25	JF320746	6	Loihi Seamount	This study
6	Loh OTU5 clone 26	FJ001799	3	Loihi Seamount	Rassa *et al.*, 2009
6	Loh OTU5 clone 49	JF320713	2	Loihi Seamount	Rassa *et al.*, 2009
7	SPL OTU1 clone 10	JF320745	8	Loihi Seamount	This study
8	2-WB OTU10 clone 8	EU574670	5	Southern Mariana Trough	Davis and Moyer, 2008
8	AV19F clone 30b	FJ905632	2	Tonga Arc	Forget *et al.*, 2010
8	ELSC clone 40	FJ205311	1	East Lau Spreading Center	GenBank; C. Dong, pers. comm.

(*Continued on next page*)

TABLE S2. Operational Taxonomic Unit (OTU) designations for the full-length *Zetaproteobacteria* dataset. (*Continued*)

OTU	Clone	Accession Number	Rep. no. clones	Region	Reference
9	Environmental enrichment JMM_S4-B-H2a	HQ206658	1	Maine	McBeth *et al.*, 2011
9	Cleft Mound pushcore clone CMB-6	DQ832644	3	Juan de Fuca Ridge	Davis *et al.*, 2009
9	Papm3 clone BL26	AB284832	1	Southern Mariana Trough	Kato *et al.*, 2009b
9	Papm3 clone BL54	AB284833	1	Southern Mariana Trough	Kato *et al.*, 2009b
9	Vailulu'u Seamount clone VS_CL-111	FJ497362	1	Vailulu'u Seamount	Sudek *et al.*, 2009
9	Vailulu'u Seamount clone VS_CL-407	FJ497659	1	Vailulu'u Seamount	Sudek *et al.*, 2009
10	Poh OTU10 clone 1	JF320727	2	Loihi Seamount	This study
10	Poh OTU11 clone 68	JF320728	2	Loihi Seamount	This study
10	ULoh OTU22 clone 69	JF320771	2	Loihi Seamount	This study
10	UNH OTU12 clone 27	JF320784	2	Loihi Seamount	This study
11	*Mariprofundus ferrooxydans* strain JV-1	EF493244	1	Loihi Seamount	Emerson and Moyer, 2002; Emerson *et al.*, 2007
11	*Mariprofundus ferrooxydans* strain PV-1	EF493243	1	Loihi Seamount	Emerson and Moyer, 2002; Emerson *et al.*, 2007
11	*Mariprofundus* sp. strain M34	JF317957	1	Loihi Seamount	This study
11	Loh OTU7 clone 5	FJ001801	2	Loihi Seamount	Rassa *et al.*, 2009
11	SFB salt marsh sediment clone WSMO200	GU291335	2	California	Moreau *et al.*, 2010
12	UNB OTU6 clone 31	JF261517	3	Loihi Seamount	Edwards *et al.*, 2011
12	UNB OTU8 clone 7	JF261519	2	Loihi Seamount	Edwards *et al.*, 2011
13	V1F clone 74b	FJ905724	5	Tonga Arc	Forget *et al.*, 2010
14	Red Sea bacterium KT-2K34	AJ309526	1	Red Sea	Eder *et al.*, 2001
14	UNB OTU7 clone 44	JF261518	3	Loihi Seamount	Edwards *et al.*, 2011
15	Papm3 clone BL17	AB284830	1	Southern Mariana Trough	Kato *et al.*, 2009b
15	Papm3 clone BL58	AB284834	1	Southern Mariana Trough	Kato *et al.*, 2009b
15	Papm3 clone BL23	AB284831	1	Southern Mariana Trough	Kato *et al.*, 2009b
16	AV19F clone 42b	FJ905640	2	Tonga Arc	Forget *et al.*, 2010
16	V1F clone 105b	FJ905738	1	Tonga Arc	Forget *et al.*, 2010
17	ELSC clone 100	FJ205312	3	East Lau Spreading Center	GenBank; C. Dong, pers. comm.
18	Laboratory enrichment JMM_Dock-D2b-C6	HQ206656	1	Maine	McBeth *et al.*, 2011
18	Laboratory enrichment JMM_S1-C-H1a	HQ206657	1	Maine	McBeth *et al.*, 2011
19	SPL OTU11 clone 31	JF320755	2	Loihi Seamount	This study
20	2-WB OTU8 clone 7	EU574668	2	Southern Mariana Trough	Davis and Moyer, 2008
21	AV19F clone 13b	FJ905621	1	Tonga Arc	Forget *et al.*, 2010
21	V1F clone 125b	FJ905748	1	Tonga Arc	Forget *et al.*, 2010
22	PV-549_X2 clone P9X2b7H12	EU491223	1	Loihi Seamount	Santelli *et al.*, 2008
22	PV-549_X2 clone P9X2b8F02	EU491311	1	Loihi Seamount	Santelli *et al.*, 2008
23	*Mariprofundus* sp. strain GSB2	HQ206653	1	Maine	McBeth *et al.*, 2011
24	Guaymas Core B clone B03R022	AY197408	1	Guaymas	Dhillon *et al.*, 2003
25	AV19F clone 106b	FJ905673	1	Tonga Arc	Forget *et al.*, 2010
26	Kermadec Arc clone CF-30	FJ535293	1	Kermadec Arc	Hodges and Olson, 2009
27	Kermadec Arc clone TS-20	FJ535342	1	Kermadec Arc	Hodges and Olson, 2009
28	Vailulu'u Seamount clone VS_CL-152	FJ497401	1	Vailulu'u Seamount	Sudek *et al.*, 2009

TABLE S3. Clone library environmental data, including pH and geochemical concentrations and ratios.

Year Collected	Clone Library	Region	pH	Fe (μM)	Mn (μM)	Fe/Mn (mol ratio)	Si (μM)	Reference
Full-Length Sequences								
1991	Pele's Vents Bacteria (PVB)	Loihi Seamount	4.2-5.5†	603-1460†	20.6-48.4†	~47.9†	~667-1770†	Karl *et al.*, 1988; Sedwick *et al.*, 1992, 1994
1996	*M. ferrooxydans* strain PV-1	Loihi Seamount	6.5-7.6	270-664	4.6-11.9	55.9-59.1	661-1480	Wheat *et al.*, 2000
1998	*M. ferrooxydans* strain JV-1	Loihi Seamount	5.3	67	59.1	1.1	2490	Wheat *et al.*, 2000
2003	PV-549_X2	Loihi Seamount						
2004	PV-601_b18 (UHO)	Loihi Seamount	5.9-6.0	261-452	11.0-19.1	22.9-23.7	2206-3810	Wheat *et al.*, 2000
2004	PV-602_b14 (SPL)	Loihi Seamount	6.1	476	8.9	53.2	3892	Wheat *et al.*, 2000
2006	*Mariprofundus* sp. strain M34	Loihi Seamount	6.0	610	17.0	35.9	4226	Glazer and Rouxel, 2009
2006	Growth Chamber LoBT_24 (Loh)	Loihi Seamount	5.9	235	21.8	10.8	1558	Glazer and Rouxel, 2009
2006	Ula Nui Bacteria (UNB)	Loihi Seamount						
2007	J2-308_redgreen (UNH)	Loihi Seamount	6.7-6.8	576-578	24.1-24.5	23.6-23.9	4069-4131	Glazer and Rouxel, 2009
2007	J2-310_bluered (ULoh)	Loihi Seamount						
2008	J2-373_scoop1 (Poh)	Loihi Seamount	5.6-6.5	507-773	12.4-19.0	40.7-40.9	1553-2275	Glazer and Rouxel, 2009
1997	Red Sea KT-2	Red Sea	6.5					Eder *et al.*, 2001
1998	Guaymas Core B	Southern Guaymas vent field						
2002	Cleft Mound pushcore 23	Juan de Fuca Ridge	7.5	0.94	9.7	0.097	241	Davis *et al.*, 2009
2003	WSMO200	California	1.8					Moreau *et al.*, 2010
2003	1-WB	Southern Mariana Trough						
2003	2-WB	Southern Mariana Trough						
2004	Papm3	Southern Mariana Trough	7.3					Kato *et al.*, 2009b
2005	Tangaroa Floc (TF)	Kermadec Arc	~5	6				Hodges and Olson, 2009
2005	Tangaroa Sediment (TS)	Kermadec Arc	~5	6				Hodges and Olson, 2009
2005	Clark Floc (CF)	Kermadec Arc	~5	15				Hodges and Olson, 2009
2005	Vailulu'u Seamount (VS_CL)	Vailulu'u Seamount	7.5-8.4					Sudek *et al.*, 2009
2007	East Lau Spreading Center (ELSC)	ELSC						
2007	R1053 (V1F)	Tonga Arc						
2007	R1046 (AV19F)	Tonga Arc						
2008	*Mariprofundus* sp. strain GSB2	Maine	6.2					McBeth *et al.*, 2011
2009/10	Bigelow Enrichment Experiments	Maine						

† Averaged from samples from 1987-1992; data not used in analyses.

SUPPLEMENTARY REFERENCES

1. **Davis, R. E., and C. L. Moyer.** 2008. Extreme spatial and temporal variability of hydrothermal microbial mat communities along the Mariana Island Arc and southern Mariana back-arc system. J. Geophys. Res. **113:**B08S15. doi:10.1029/2007JB005413.

2. **Davis, R. E., D. S. Stakes, C. G. Wheat, and C. L. Moyer.** 2009. Bacterial variability within an iron-silica-manganese-rich hydrothermal mound located off-axis at the Cleft Segment, Juan de Fuca Ridge. Geomicrobiol. J. **26:**570-580.

3. **Dhillon, A., A. Teske, J. Dillon, D. A. Stahl, and M. L. Sogin.** 2003. Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. Appl. Environ. Microbiol. **69:**2765-2772.

4. **Eder, W., L. L. Jahnke, M. Schmidt, and R. Huber.** 2001. Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. Appl. Environ. Microbiol. **67:**3077-3085.

5. **Edwards, K. J., B. T. Glazer, O. J. Rouxel, W. Bach, D. Emerson, R. E. Davis, B. M. Toner, C. S. Chan, B. M. Tebo, H. Staudigel, and C. L. Moyer.** May 2011, posting date. Ultra-diffuse hydrothermal venting supports Fe-oxidizing bacteria and massive umber deposition at 5000 m off Hawaii. ISME J. doi:10.1038/ismej.2011.48.

6. **Emerson, D., and C. L. Moyer.** 2002. Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition. Appl. Environ. Microbiol. **68:**3085-3093.

7. **Emerson, D., J. A. Rentz, T. G. Lilburn, R. E. Davis, H. Aldrich, C. Chan, and C. L. Moyer.** 2007. A novel lineage of *Proteobacteria* involved in formation of marine Fe-

oxidizing microbial mat communities. PLoS ONE **2:**e667. doi:10.1371/journal.pone.0000667.

8. **Forget, N. L., S. A. Murdock, and S. K. Juniper.** 2010. Bacterial diversity in Fe-rich hydrothermal sediments at two South Tonga Arc submarine volcanoes. Geobiology **8:**417-432. doi:10.1111/j.1472-4669.2010.00247.x.

9. **Glazer, B. T., and O. J. Rouxel.** 2009. Redox speciation and distribution within diverse iron-dominated microbial habitats at Loihi Seamount. Geomicrobiol. J. **26:**606-622.

10. **Hodges, T. W., and J. B. Olson.** 2009. Molecular comparison of bacterial communities within iron-containing flocculent mats associated with submarine volcanoes along the Kermadec Arc. Appl. Environ. Microbiol. **75:**1650-1657.

11. **Karl, D. M., G. M. McMurtry, A. Malahoff, and M. O. Garcia.** 1988. Loihi Seamount, Hawaii: a mid-plate volcano with a distinctive hydrothermal system. Nature **335:**532-535.

12. **Kato, S., K. Yanagawa, M. Sunamura, Y. Takano, J. Ishibashi, T. Kakegawa, M. Utsumi, T. Yamanaka, T. Toki, T. Noguchi, K. Kobayashi, A. Moroi, H. Kimura, Y. Kawarabayasi, K. Marumo, T. Urabe, and A. Yamagishi.** 2009b. Abundance of *Zetaproteobacteria* within crustal fluids in back-arc hydrothermal fields of the Southern Mariana Trough. Environ. Microbiol. **11:**3210-3222.

13. **McBeth, J. M., B. J. Little, R. I. Ray, K. M. Farrar, and D. Emerson.** 2011. Neutrophilic iron-oxidizing "*Zetaproteobacteria*" and mild steel corrosion in nearshore marine environments. Appl. Environ. Microbiol. **77:**1405-1412.

14. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1995. Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **61:**1555-1562.

15. **Rassa, A. C., S. M. McAllister, S. A. Safran, and C. L. Moyer.** 2009. *Zeta-Proteobacteria* dominate the colonization and formation of microbial mats in low-temperature hydrothermal vents at Loihi Seamount, Hawaii. Geomicrobiol. J. **26:**623-638.

16. **Santelli, C. M., B. N. Orcutt, E. Banning, W. Bach, C. L. Moyer, M. L. Sogin, H. Staudigel, and K. J. Edwards.** 2008. Abundance and diversity of microbial life in ocean crust. Nature **453:**653-656.

17. **Sedwick, P. N., G. M. McMurtry, D. R. Hilton, and F. Goff.** 1994. Carbon dioxide and helium in hydrothermal fluids from Loihi Seamount, Hawaii, USA: temporal variability and implications for the release of mantle volatiles. Geochim. Cosmochim. Acta **58:**1219-1227.

18. **Sedwick, P. N., G. M. McMurtry, and J. D. Macdougall.** 1992. Chemistry of hydrothermal solutions from Pele's Vents, Loihi Seamount, Hawaii. Geochim. Cosmochim. Acta **56:**3643-3667.

19. **Sudek, L. A., A. S. Templeton, B. M. Tebo, and H. Staudigel.** 2009. Microbial ecology of Fe (hydr)oxide mats and basaltic rock from Vailulu'u Seamount, American Samoa. Geomicrobiol. J. **26:**581-596.

20. **Wheat, C. G., H. W. Jannasch, J. N. Plant, C. L. Moyer, F. J. Sansone, and G. M. McMurtry.** 2000. Continuous sampling of hydrothermal fluids from Loihi Seamount after the 1996 event. J. Geophys. Res. **105:**19353-19367.

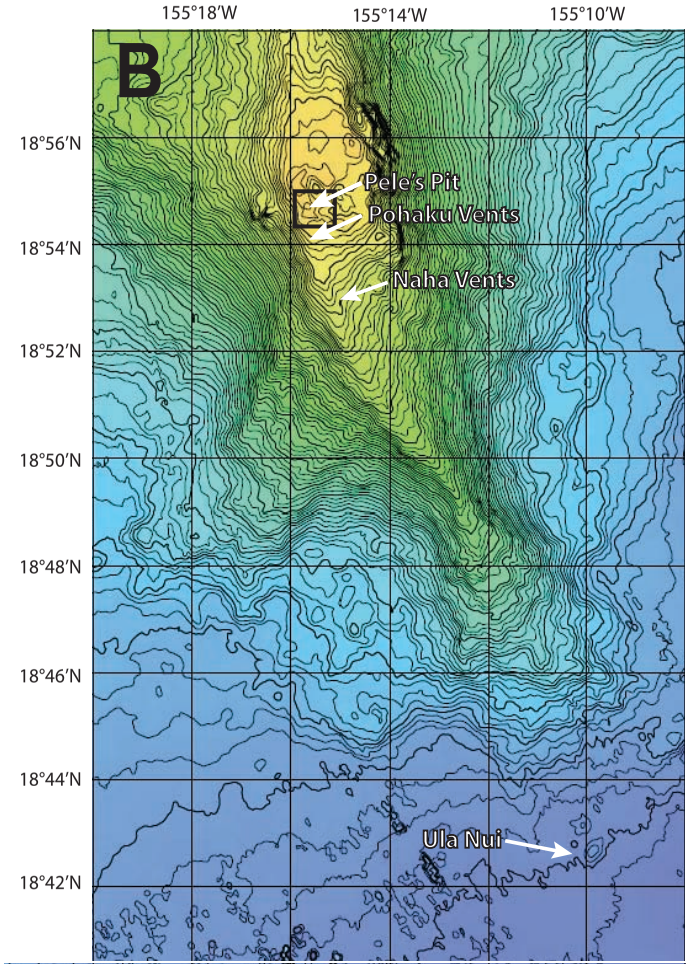
SUMMARY OF APPENDICES

APPENDIX A: Additional information for the five novel clone libraries

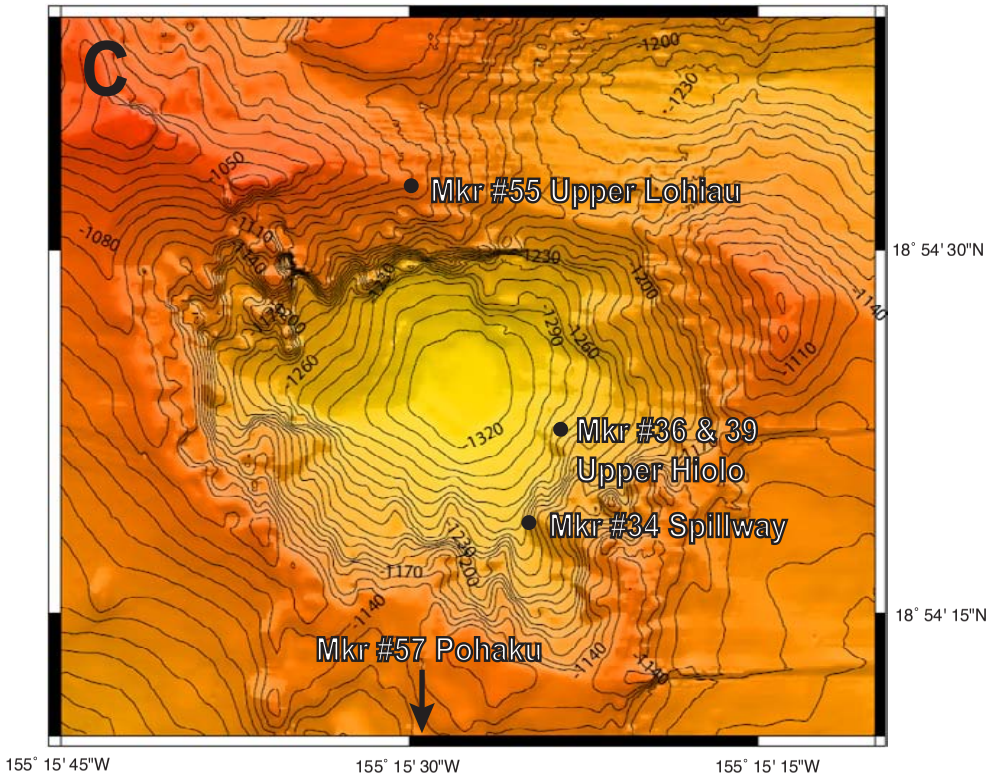
APPENDIX B: Additional information regarding *Zetaproteobacteria* sequences

APPENDIX A

FIG. A1. Maps of Loihi Seamount, Hawaii showing increasing detail of sampling area. A) Map of Hawaii with the location of Loihi Seamount indicated by a star. B) Bathymetric map of Loihi Seamount showing the locations of sites of interest for this study, including Pele's Pit, Pohaku Vents, and Ula Nui (FeMO Deep). C) Bathymetric map of Pele's Pit [detail of area indicated in map B] showing the sampling sites for the five clone libraries constructed for this study. Maps B and C modified from Rassa *et al.* (2009).


B
155°18'W
155°14'W
155°10'W
18°56'N
18°54'N
18°52'N
18°50'N
18°48'N
18°46'N
18°44'N
18°42'N
Pele's Pit
Pohaku Vents
Naha Vents
Ula Nui


Hawaii
A


C
Mkr #55 Upper Lohiau
Mkr #36 & 39
Upper Hiolo
Mkr #34 Spillway
Mkr #57 Pohaku
18° 54' 30"N
18° 54' 15"N
155° 15' 45"W
155° 15' 30"W
155° 15' 15"W

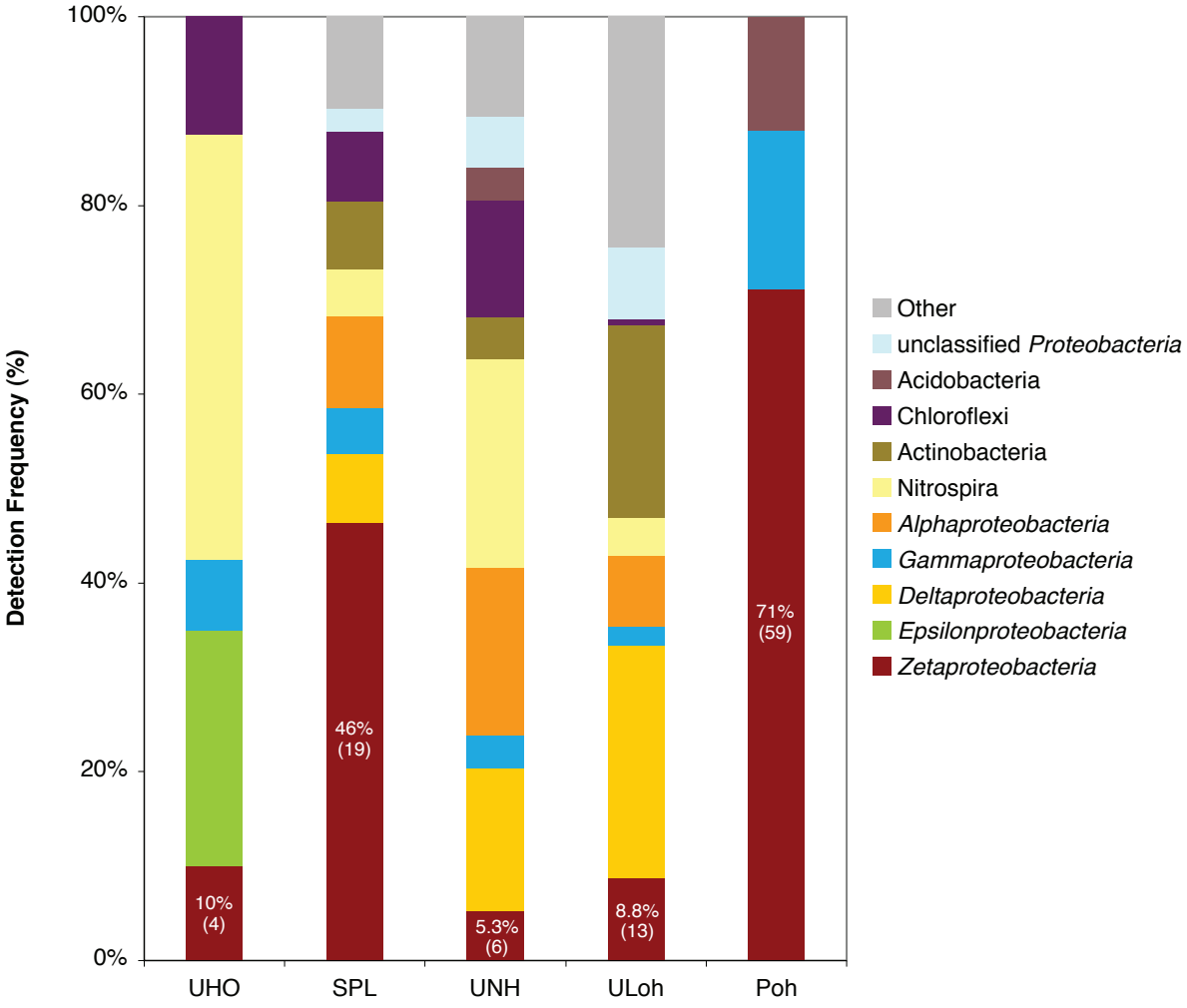


FIG. A2. Stacked bar graph showing class- and phylum-level phylogenetic affiliations for the clones from all five of the novel clone libraries constructed for this study. Percentages are shown for the *Zetaproteobacteria*, with total number of clones in parentheses.

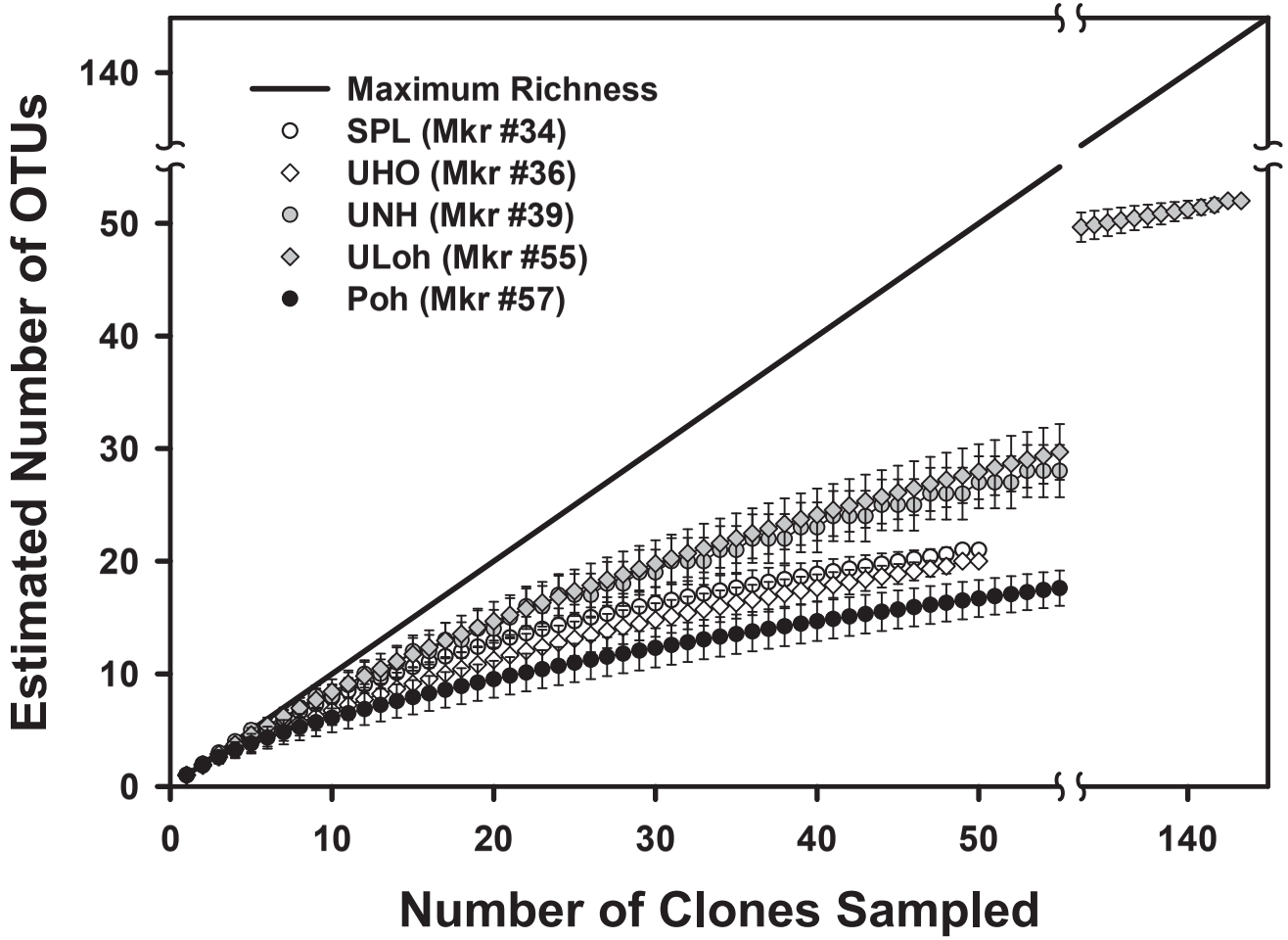


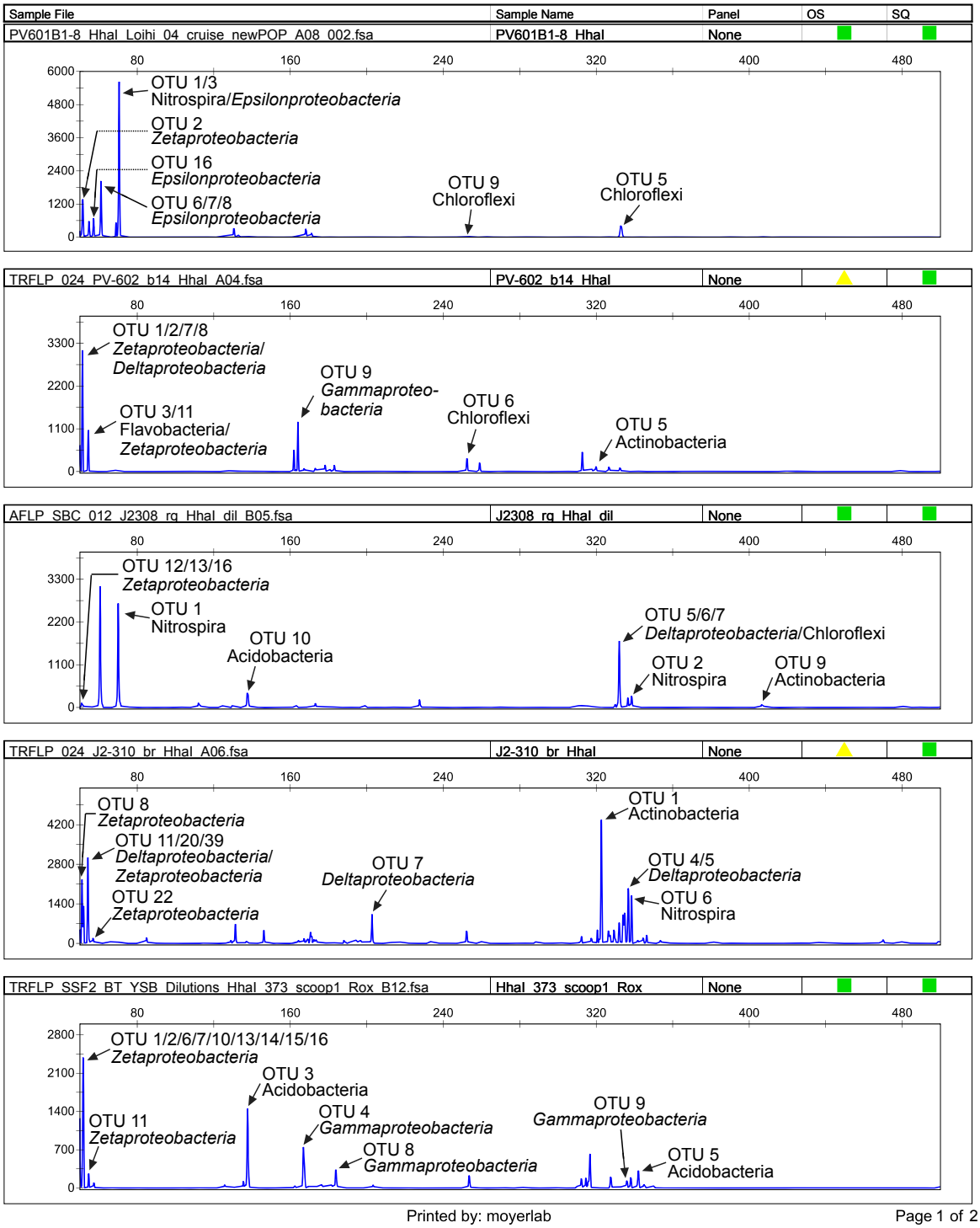
FIG. A3. Rarefaction curves comparing the estimated population richness between the five clone libraries constructed for this study. OTU richness is not significantly different between the UNH & ULoh and SPL & UHO libraries, whereas UNH/ULoh, SPL/UHO, and Poh are significantly different from each other. Error bars represent standard deviation.

FIG. A4. Terminal-restriction fragment length polymorphism (T-RFLP) electropherograms showing *Hha*I (A) and *Rsa*I (B) digests for the five samples from which clone libraries were constructed for this study. T-RFLP is a molecular community fingerprinting technique (see Rassa *et al.*, 2009). Samples are (from top to bottom): PV-601_b18 (UHO), PV-602_b14 (SPL), J2-308_redgreen (UNH), J2-310_bluered (ULoh), and J2-373_scoop1 (Poh). Arrows indicate peaks corresponding to OTUs detected in each clone libraries with associated phylogenetic affiliation indicated.

Applied Biosystems
GeneMapper 4.0

A) *Hha*I

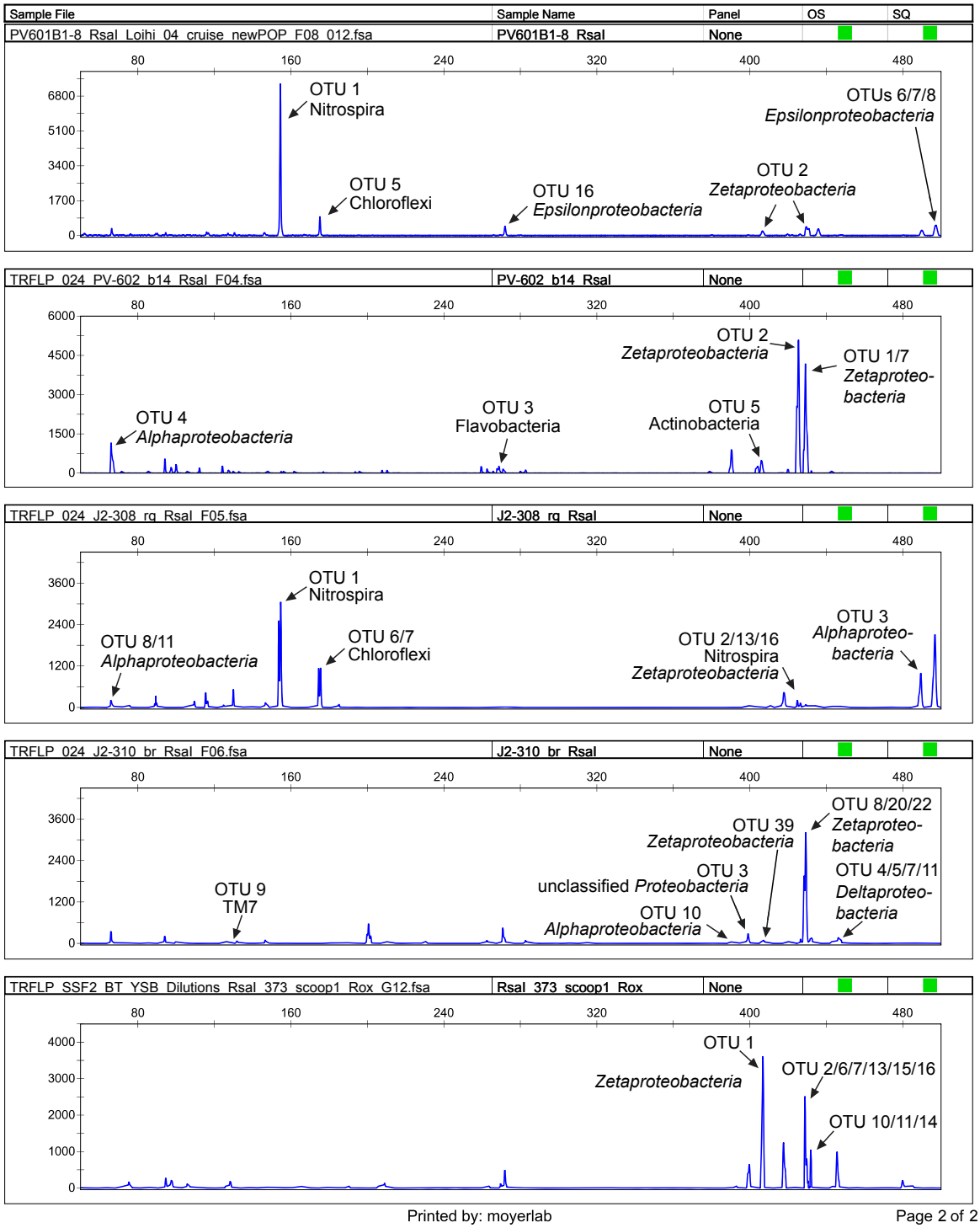
Novel Clone Libraries



Applied Biosystems
GeneMapper 4.0

B) *RsaI*

Novel Clone Libraries



Printed by: moyerlab

Page 2 of 2

TABLE A1. OTU determination, phylogenetic grouping, closest cultured representative with associated S_{ab} and similarity scores, predicted physiology, and oxygen requirements for the non-*Zetaproteobacteria* OTUs from the five novel clone libraries.

Clone Library	OTU	Sequenced clone no.	No. of clones (% of library)	Phylogenetic Grouping[1]	Closest Cultured Representative[1]	S_{ab} Score [1]	Similarity (%)[1]	Predicted Physiology Type[2]	Oxygen Requirement
PV-601_b18 (UHO)	1	6 & 16	18 (36.0)	Nitrospira	*Thermodesulfovibrio* sp.	0.577	89.4	SRB, HOB	Anaerobic
	3	150	3 (6.0)	ε–*Proteobacteria*	*Hydrogenimonas thermophila*	0.883	98.1	HOB, SRB, NRB	Microaerophilic to Anaerobic
	4	158	3 (6.0)	γ–*Proteobacteria*	*Thiomicrospira* sp.	0.720	94.7	SOB	Aerobic
	5	202	3 (6.0)	Chloroflexi	n.a.	n.a.	n.a.	n.a.	n.a.
	6	209	2 (4.0)	ε–*Proteobacteria*	*Nitratiruptor* sp.	0.751	95.2	NRB, HOB	Microaerophilic
	7	65	2 (4.0)	ε–*Proteobacteria*	*Nitratiruptor* sp.	0.758	95.4	NRB, HOB	Microaerophilic
	8	125	2 (4.0)	ε–*Proteobacteria*	*Nitratiruptor* sp.	0.750	95.2	NRB, HOB	Microaerophilic
	9	34	2 (4.0)	Chloroflexi	n.a.	n.a.	n.a.	n.a.	n.a.
	16	139	1 (2.0)	ε–*Proteobacteria*	*Sulfurimonas autotrophica*	0.936	98.4	SOB	Aerobic
PV-602_b14 (SPL)	3	216	4 (8.0)	Flavobacteria	*Actibacter* sp.	0.678	92.7	Heterotroph	Aerobic
	4	116	4 (8.0)	α–*Proteobacteria*	*Roseobacter* sp.	0.762	94.7	Heterotroph	Aerobic
	5	40	3 (6.0)	Actinobacteria	*Iamia* sp.	0.683	93.9	n.a.	n.a.
	6	17	3 (6.0)	Chloroflexi	n.a.	n.a.	n.a.	n.a.	n.a.
	8	7	3 (6.0)	δ–*Proteobacteria*	*Enhygromyxa* sp.	0.584	89.9	n.a.	n.a.
	9	5	2 (4.0)	γ–*Proteobacteria*	*Acinetobacter* sp.	0.988	99.8	n.a.	n.a.
	10	28	2 (4.0)	Nitrospira	*Thermodesulfovibrio* sp.	0.610	89.0	SRB, HOB	Anaerobic
	12	69	1 (2.0)	unclassified *Proteobacteria*	n.a.	n.a.	n.a.	n.a.	n.a.
J2-308_redgreen (UNH)	1	113	14 (12.4)	Nitrospira	*Thermodesulfovibrio* sp.	0.610	89.3	SRB, HOB	Anaerobic
	2	6	10 (8.8)	unclassified Nitrospira	*Thermodesulfovibrio* sp.	0.541	88.1	SRB, HOB	Anaerobic
	3	77	8 (7.1)	α–*Proteobacteria*	*Thalassobius* sp.	0.775	95.1	Heterotroph	Aerobic
	4	22	6 (5.3)	unclassified *Proteobacteria*	n.a.	n.a.	n.a.	n.a.	n.a.
	5	136	6 (5.3)	δ–*Proteobacteria*	n.a.	n.a.	n.a.	n.a.	n.a.
	6	30	6 (5.3)	Chloroflexi	*Dehalococcoides* sp.	0.525	88.7	n.a.	n.a.
	7	34	6 (5.3)	Chloroflexi	*Dehalococcoides* sp.	0.525	88.7	n.a.	n.a.
	8	110	5 (4.4)	α–*Proteobacteria*	*Roseobacter* sp.	0.763	94.7	Heterotroph	Aerobic
	9	141	5 (4.4)	Actinobacteria	*Rhodococcus erythropolis*	0.985	99.7	n.a.	n.a.
	10	75	4 (3.5)	Acidobacteria	n.a.	n.a.	n.a.	n.a.	n.a.
	11	114	4 (3.5)	α–*Proteobacteria*	*Sulfitobacter* sp.	0.722	93.3	SOB, Heterotroph	Aerobic
	14	118	2 (1.8)	WS3	n.a.	n.a.	n.a.	n.a.	n.a.
J2-310_bluered (ULoh)	1	110	23 (15.6)	Actinobacteria	*Conexibacter* sp.	0.561	87.5	n.a.	n.a.
	2	159	9 (6.1)	WS3	n.a.	n.a.	n.a.	n.a.	n.a.
	3	87	9 (6.1)	unclassified *Proteobacteria*	n.a.	n.a.	n.a.	n.a.	n.a.
	4	160	8 (5.4)	δ–*Proteobacteria*	*Desulforhabdus* sp.	0.540	87.4	SRB, HOB; Heterotroph	Anaerobic
	5	48	7 (4.8)	δ–*Proteobacteria*	*Desulforhabdus* sp.	0.541	87.5	SRB, HOB; Heterotroph	Anaerobic
	6	161	6 (4.1)	unclassified Nitrospira	*Thermodesulfovibrio* sp.	0.542	88.1	SRB, HOB	Anaerobic
	7	177	6 (4.1)	δ–*Proteobacteria*	n.a.	n.a.	n.a.	n.a.	n.a.
	9	187	5 (3.4)	TM7	n.a.	n.a.	n.a.	n.a.	n.a.
	10	123	5 (3.4)	α–*Proteobacteria*	*Thalassobius* sp.	0.804	95.4	Heterotroph	Aerobic
	11	68	4 (2.7)	δ–*Proteobacteria*	*Pelobacter* sp.	0.543	88.2	Heterotroph	Anaerobic
	12	2	3 (2.0)	Actinobacteria	*Rhodococcus erythropolis*	0.990	99.8	n.a.	n.a.
	13	84	3 (2.0)	unclassified Bacteria	n.a.	n.a.	n.a.	n.a.	n.a.
J2-373_scoop1 (Poh)	3	24	7 (8.3)	Acidobacteria	n.a.	n.a.	n.a.	n.a.	n.a.
	4	76	4 (4.8)	γ–*Proteobacteria*	*Methylobacter* sp.	0.754	95.4	Methanotroph	Aerobic
	5	48	3 (3.6)	Acidobacteria	n.a.	n.a.	n.a.	n.a.	n.a.
	8	10	2 (2.4)	γ–*Proteobacteria*	*Methylobacter* sp.	0.720	95.0	Methanotroph	Aerobic
	9	82	2 (2.4)	γ–*Proteobacteria*	*Marinicella* sp.	0.651	92.1	n.a.	n.a.

[1]As determined by RDP seqmatch release 10.26.

[2]SRB = Sulfur-Reducing Bacteria; HOB = Hydrogen-Oxidizing Bacteria; NRB = Nitrate-Reducing Bacteria; SOB = Sulfur-Oxidizing Bacteria

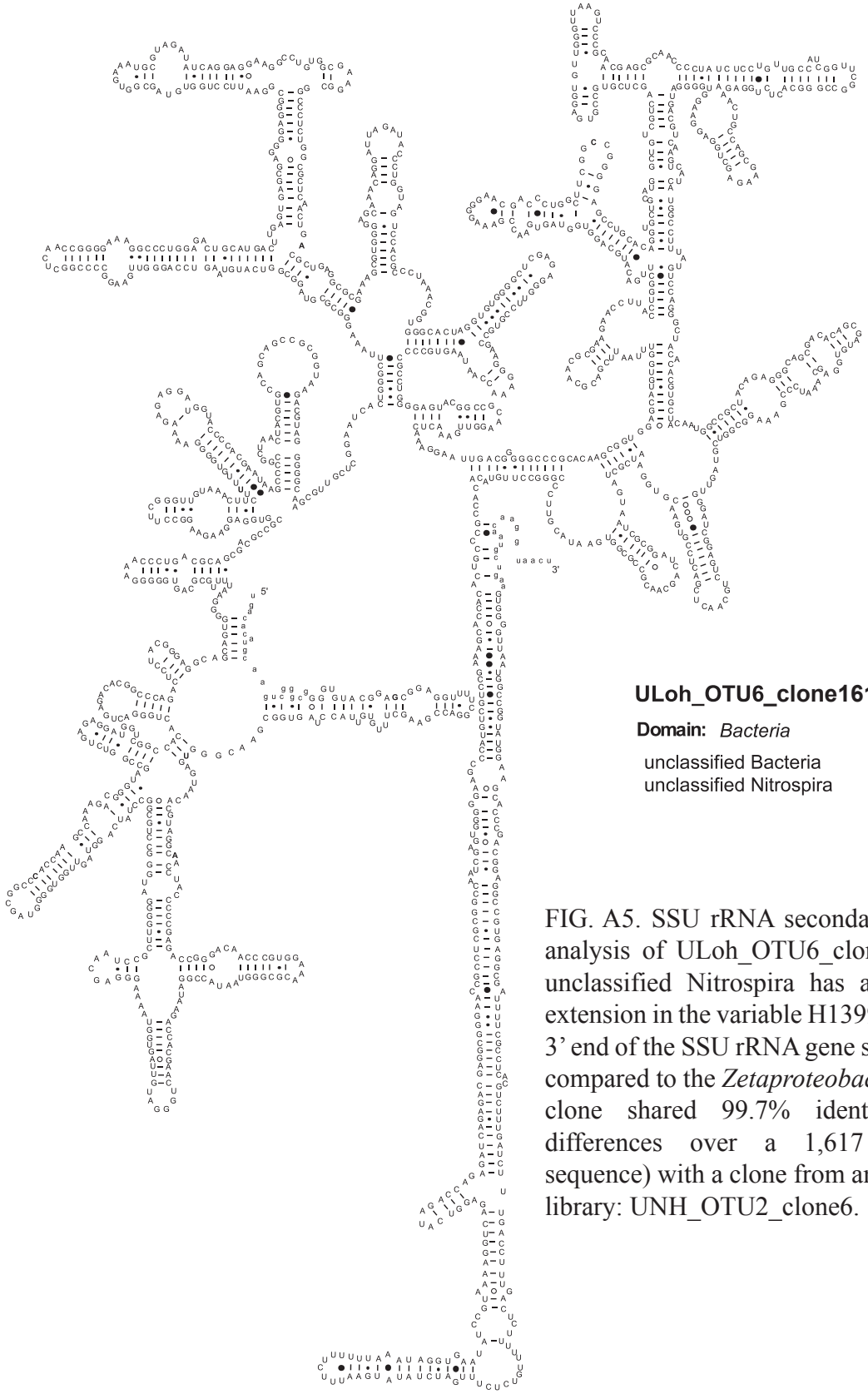


FIG. A5. SSU rRNA secondary structure analysis of ULoh_OTU6_clone161. This unclassified Nitrospira has an ~150 bp extension in the variable H1399 loop at the 3' end of the SSU rRNA gene sequence (as compared to the *Zetaproteobacteria*). This clone shared 99.7% identity (5 bp differences over a 1,617 bp gene sequence) with a clone from another clone library: UNH_OTU2_clone6.

APPENDIX B

Below is an exhaustive list of locations where *Zetaproteobacteria* have been detected, including sites with only partial-length sequences not used in this study:

- Microbial mats, altered Fe-oxide-stained basalts, and microbial growth chamber experiments at Loihi Seamount (Moyer *et al.*, 1995; Emerson and Moyer, 2002; Emerson *et al.*, 2007; Santelli *et al.*, 2008; Rassa *et al.*, 2009; Edwards *et al.*, 2011)
- Microbial mats and borehole fluids at the Southern Mariana Trough (Davis and Moyer, 2008; Kato *et al.*, 2009a; Kato *et al.*, 2009b)
- Fe-oxide hydrothermal sediments and chimneys at the Tonga Arc (Forget *et al.*, 2010)
- Hydrothermal sediments at the East Lau Spreading Center (Dong and Shao, Genbank FJ205309-FJ205312)
- Fe-flocculent mats and sediments along the Kermadec Arc (Hodges and Olson, 2009)
- Microbial mat and basalt samples from Vailulu'u Seamount (Sudek *et al.*, 2009)
- Iron-silica-manganese-rich hydrothermal mound sediments from off-axis Cleft Segment, Juan de Fuca Ridge (Davis *et al.*, 2009)
- Hydrothermal sediments from the Santorini flooded caldera, Greece (Handley *et al.*, 2010)
- Geothermal springs at Edipos hot springs, Greece (Kormas *et al.*, 2009)
- Hydrothermal sediments in the Guaymas Basin (Dhillon *et al.*, 2003)
- Acid-mine-drainage-impacted salt marsh sediments in San Francisco Bay (Moreau *et al.*, 2010)
- Mild steel corrosion enrichment experiments conducted in near-shore marine and salt marsh environments, Maine (McBeth *et al.*, 2011)
- Brine-seawater interface at Kebrit Deep, Red Sea (Eder *et al.*, 2001)
- Mid-Atlantic Ridge *Rimicaris exoculata* gut (Zbinden and Cambon-Bonavita, 2003)
- Antarctica marine continental shelf sediment (Bowman and McCuaig, 2003)

ADDITIONAL REFERENCES

1. **Bowman, J. P., and R. D. McCuaig.** 2003. Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. Appl. Environ. Microbiol. **69:**2463-2483.

2. **Kormas, K. A., H. Tamaki, S. Hanada, and Y. Kamagata.** 2009. Apparent richness and community composition of Bacteria and Archaea in geothermal springs. Aquat. Microb. Ecol. **57:**113-122.

3. **Zbinden, M., and M. Cambon-Bonavita.** 2003. Occurrence of *Deferribacterales* and *Entomoplasmatales* in the deep-sea Alvinocarid shrimp *Rimicaris exoculata* gut. FEMS Microbiol. Ecol. **46:**23-30.

TABLE B1. OTU designations, seqmatch (S_{ab}) scores, and similarity scores for the full-length *Zetaproteobacteria* dataset.

OTU	Clone	Accession Number	Rep. no. clones	Region	Sab Score[1]	Similarity (%)[1]
1	Loh OTU2 clone 60	FJ001796	11	Loihi Seamount	0.823	96.7
1	Poh OTU1 clone 9	JF320714	8	Loihi Seamount	0.811	96.0
1	Poh OTU1 clone 67	JF320715	8	Loihi Seamount	0.815	96.0
1	Poh OTU1 clone 74	JF320716	8	Loihi Seamount	0.818	96.2
1	Poh OTU15 clone 52	JF320731	1	Loihi Seamount	0.826	96.3
1	ULoh OTU20 clone 55	JF320770	2	Loihi Seamount	0.819	96.4
1	ULoh OTU39 clone 94	JF320772	1	Loihi Seamount	0.831	96.2
1	UNH OTU16 clone 122	JF320787	2	Loihi Seamount	0.823	96.4
1	V1F clone 151b	FJ905756	6	Tonga Arc	0.850	96.8
1	AV19F clone 4b	FJ905617	17	Tonga Arc	0.803	96.4
1	ELSC clone 13	FJ205309	1	East Lau Spreading Center	0.835	96.6
1	Kermadec Arc clone TF-31	FJ535254	1	Kermadec Arc	0.805	95.7
1	1-WB OTU4 clone 29	EU574657	1	Southern Mariana Trough	0.820	96.6
2	PVB OTU4 clone 13	U15116	1	Loihi Seamount	0.725	94.6
2	UHO OTU2 clone 21	JF320735	4	Loihi Seamount	0.734	94.6
2	SPL OTU7 clone 22	JF320751	3	Loihi Seamount	0.749	95.3
2	Loh OTU1 clone 67	FJ001795	12	Loihi Seamount	0.736	94.6
2	Poh OTU2 clone 12	JF320717	7	Loihi Seamount	0.745	94.4
2	Poh OTU2 clone 27	JF320718	7	Loihi Seamount	0.728	94.7
2	Poh OTU2 clone 89	JF320719	7	Loihi Seamount	0.737	94.7
2	ULoh OTU8 clone 191	JF320764	8	Loihi Seamount	0.725	94.4
2	UNH OTU13 clone 95	JF320785	2	Loihi Seamount	0.731	94.6
2	Cleft Mound pushcore clone CMB-2	DQ832638	2	Juan de Fuca Ridge	0.722	94.6
2	Vailulu'u Seamount clone VS_CL-318	FJ497570	1	Vailulu'u Seamount	0.699	94.7
3	Poh OTU16 clone 64	JF320732	1	Loihi Seamount	0.811	96.3
3	Cleft Mound pushcore clone CMB-25	DQ832637	3	Juan de Fuca Ridge	0.828	96.3
3	AV19F clone 45b	FJ905642	5	Tonga Arc	0.842	96.8
3	V1F clone 2b	FJ905692	27	Tonga Arc	0.845	96.8
4	Poh OTU6 clone 78	JF320723	3	Loihi Seamount	0.656	92.8
4	Poh OTU7 clone 34	JF320724	3	Loihi Seamount	0.690	93.0
4	Poh OTU13 clone 3	JF320729	1	Loihi Seamount	0.688	92.7
4	Poh OTU14 clone 5	JF320730	1	Loihi Seamount	0.647	92.0
4	V1F clone 118b	FJ905745	1	Tonga Arc	0.659	92.9
4	V1F clone 7b	FJ905694	1	Tonga Arc	0.663	92.7
4	V1F clone 25b	FJ905704	9	Tonga Arc	0.699	93.3
4	ELSC clone 16	FJ205310	15	East Lau Spreading Center	0.687	92.7
5	V1F clone 48b	FJ905712	12	Tonga Arc	0.666	92.8
6	SPL OTU2 clone 25	JF320746	6	Loihi Seamount	0.747	95.3
6	Loh OTU5 clone 26	FJ001799	3	Loihi Seamount	0.761	95.3
6	Loh OTU5 clone 49	JF320713	2	Loihi Seamount	0.758	95.5
7	SPL OTU1 clone 10	JF320745	8	Loihi Seamount	0.720	93.9
8	2-WB OTU10 clone 8	EU574670	5	Southern Mariana Trough	0.749	95.0
8	AV19F clone 30b	FJ905632	2	Tonga Arc	0.771	95.5
8	ELSC clone 40	FJ205311	1	East Lau Spreading Center	0.764	95.4

(*Continued on next page*)

TABLE B1. OTU designations, seqmatch (S_{ab}) scores, and similarity scores for the full-length *Zetaproteobacteria* dataset. (*Continued*)

OTU	Clone	Accession Number	Rep. no. clones	Region	Sab Score[1]	Similarity (%)[1]
9	Environmental enrichment JMM_S4-B-H2a	HQ206658	1	Maine	0.748	95.1
9	Cleft Mound pushcore clone CMB-6	DQ832644	3	Juan de Fuca Ridge	0.720	94.1
9	Papm3 clone BL26	AB284832	1	Southern Mariana Trough	0.725	94.4
9	Papm3 clone BL54	AB284833	1	Southern Mariana Trough	0.744	94.6
9	Vailulu'u Seamount clone VS_CL-111	FJ497362	1	Vailulu'u Seamount	0.739	94.7
9	Vailulu'u Seamount clone VS_CL-407	FJ497659	1	Vailulu'u Seamount	0.734	94.9
10	Poh OTU10 clone 1	JF320727	2	Loihi Seamount	0.617	90.6
10	Poh OTU11 clone 68	JF320728	2	Loihi Seamount	0.620	90.6
10	ULoh OTU22 clone 69	JF320771	2	Loihi Seamount	0.624	90.6
10	UNH OTU12 clone 27	JF320784	2	Loihi Seamount	0.622	90.5
11	*Mariprofundus ferrooxydans* strain JV-1	EF493244	1	Loihi Seamount	1.000	100.0
11	*Mariprofundus ferrooxydans* strain PV-1	EF493243	1	Loihi Seamount	1.000	100.0
11	*Mariprofundus* sp. strain M34	JF317957	1	Loihi Seamount	0.987	100.0
11	Loh OTU7 clone 5	FJ001801	2	Loihi Seamount	0.885	97.8
11	SFB salt marsh sediment clone WSMO200	GU291335	2	California	0.958	98.8
12	UNB OTU6 clone 31	JF261517	3	Loihi Seamount	0.764	95.4
12	UNB OTU8 clone 7	JF261519	2	Loihi Seamount	0.765	95.7
13	V1F clone 74b	FJ905724	5	Tonga Arc	0.662	92.7
14	Red Sea bacterium KT-2K34	AJ309526	1	Red Sea	0.826	97.0
14	UNB OTU7 clone 44	JF261518	3	Loihi Seamount	0.805	96.5
15	Papm3 clone BL17	AB284830	1	Southern Mariana Trough	0.645	91.2
15	Papm3 clone BL58	AB284834	1	Southern Mariana Trough	0.645	91.2
15	Papm3 clone BL23	AB284831	1	Southern Mariana Trough	0.646	91.3
16	AV19F clone 42b	FJ905640	2	Tonga Arc	0.594	90.7
16	V1F clone 105b	FJ905738	1	Tonga Arc	0.587	90.6
17	ELSC clone 100	FJ205312	3	East Lau Spreading Center	0.746	95.0
18	Laboratory enrichment JMM_Dock-D2b-C6	HQ206656	1	Maine	0.813	96.7
18	Laboratory enrichment JMM_S1-C-H1a	HQ206657	1	Maine	0.821	96.9
19	SPL OTU11 clone 31	JF320755	2	Loihi Seamount	0.662	92.2
20	2-WB OTU8 clone 7	EU574668	2	Southern Mariana Trough	0.717	94.1
21	AV19F clone 13b	FJ905621	1	Tonga Arc	0.798	96.3
21	V1F clone 125b	FJ905748	1	Tonga Arc	0.809	96.6
22	PV-549_X2 clone P9X2b7H12	EU491223	1	Loihi Seamount	0.591	89.8
22	PV-549_X2 clone P9X2b8F02	EU491311	1	Loihi Seamount	0.603	89.8
23	*Mariprofundus* sp. strain GSB2	HQ206653	1	Maine	0.856	97.6
24	Guaymas Core B clone B03R022	AY197408	1	Guaymas	0.706	93.3
25	AV19F clone 106b	FJ905673	1	Tonga Arc	0.724	94.5
26	Kermadec Arc clone CF-30	FJ535293	1	Kermadec Arc	0.732	94.3
27	Kermadec Arc clone TS-20	FJ535342	1	Kermadec Arc	0.702	93.3
28	Vailulu'u Seamount clone VS_CL-152	FJ497401	1	Vailulu'u Seamount	0.867	97.4

[1]As determined by RDP seqmatch release 10.26 by comparing sequences to *Mariprofundus ferrooxydans* PV-1.

TABLE B2. Analysis of Molecular Variance (AMOVA) results for all grouping strategies and sequence subsets. Significant P-values highlighted.

			Percentage of Variation			Among groups	
Sequence Subset	Grouped by:	Subgroup:	Among groups	Among clone libs within groups	Within clone libs	d.f.	P-value
All Sequences	Region	Condensed	8.44	25.16	66.40	7	**0.01743±0.00145**
		Split	8.86	24.71	66.43	9	**0.02624±0.00152**
	Temperature	3 groups	2.67	29.89	67.44	2	0.10168±0.00332
		4 groups	4.11	29.16	66.73	3	0.07020±0.00247
		5 groups	2.61	29.49	67.90	4	0.09752±0.00300
		7 groups	9.53	23.25	67.22	6	**0.00168±0.00042**
		8 groups	9.85	22.92	67.22	7	**0.00277±0.00049**
		11 groups	11.98	20.95	67.07	10	**0.00139±0.00037**
	Depth	2 groups; oxygen proxy	-1.12	32.31	68.81	1	0.44485±0.00479
		3 groups	0.04	31.56	68.40	2	0.33099±0.00443
		6 groups	7.31	25.92	66.77	5	0.05109±0.00227
	Sample Type	Full-Length Only	6.28	28.09	65.63	4	**0.04574±0.00198**
		Full & Partials	15.81	21.76	62.44	4	**0.00000±0.00000**
	OTU	n.a.	75.91	24.09	0.00	27	**0.00000±0.00000**
Main Sampling Sites Only	Region	Condensed	8.19	25.45	66.36	2	**0.01257±0.00096**
		Split	9.12	24.78	66.10	2	**0.03545±0.00191**
	Temperature	3groups; condensed	2.60	29.80	67.60	2	0.11257±0.00269
		3groups; split	2.61	29.86	67.53	2	0.13475±0.00350
		11groups; condensed	13.78	19.16	67.05	10	**0.00218±0.00044**
		11groups; split	17.14	16.26	66.60	9	**0.00059±0.00024**
	OTU	Condensed	76.62	23.38	0.00	24	**0.00000±0.00000**
		Split	76.78	23.22	0.00	21	**0.00000±0.00000**
Subset with known [Fe]	Region	n.a.	3.33	21.30	75.37	2	0.24525±0.00439
	Temperature	3 groups	4.20	19.61	76.19	2	0.16059±0.00382
		11 groups	9.77	15.13	75.10	6	0.06396±0.00225
	total Fe (μM)	n.a.	-3.47	25.54	77.93	3	0.37119±0.00463
Subset with known Fe/Mn, [Mn], [Si]	Region	n.a.	1.33	21.13	77.53	1	0.67317±0.00384
	Temperature	3 groups	4.67	18.25	77.08	2	0.21287±0.00432
	Fe/Mn	4 groups	-0.25	21.83	78.42	3	0.28970±0.00420
	[Mn]	4 groups	-7.68	28.60	79.08	3	0.68139±0.00446
	[Si]	4 groups	-7.03	28.19	78.84	3	0.63436±0.00510
Subset with known pH	Region	n.a.	10.25	19.74	70.01	5	0.05693±0.00237
	Temperature	3 groups	6.00	20.73	73.27	2	**0.04455±0.00203**
	pH	5 groups	7.55	21.12	71.33	4	0.06000±0.00242
OTU 1	Region	Condensed	12.63	82.94	4.43	2	0.11832±0.00349
		Split	15.95	79.66	4.39	3	0.12822±0.00358
	Temperature	3 groups	-11.98	106.87	5.11	1	0.93851±0.00250
		11 groups	11.83	83.69	4.48	4	0.29356±0.00486
OTU 2	Region	n.a.	29.96	30.52	39.52	2	0.10663±0.00314
	Temperature	3 groups	2.41	45.00	52.59	2	0.23238±0.00404
		11 groups	15.43	35.18	49.39	4	0.08683±0.00275
Mats Only (full-length)	Region	3 regions	8.31	19.85	71.84	2	0.02752±0.00144
		2 regions	8.64	19.45	71.91	1	0.01990±0.00118
	Temperature	3 regions	4.71	23.35	71.94	2	0.12376±0.00326
		2 regions	4.97	23.00	72.03	2	0.14386±0.00313
Mats Only (partials)	Region	n.a.	0.17	22.82	77.01	2	0.39277±0.00447
	Temperature	n.a.	-0.83	23.47	77.36	2	0.52000±0.00509

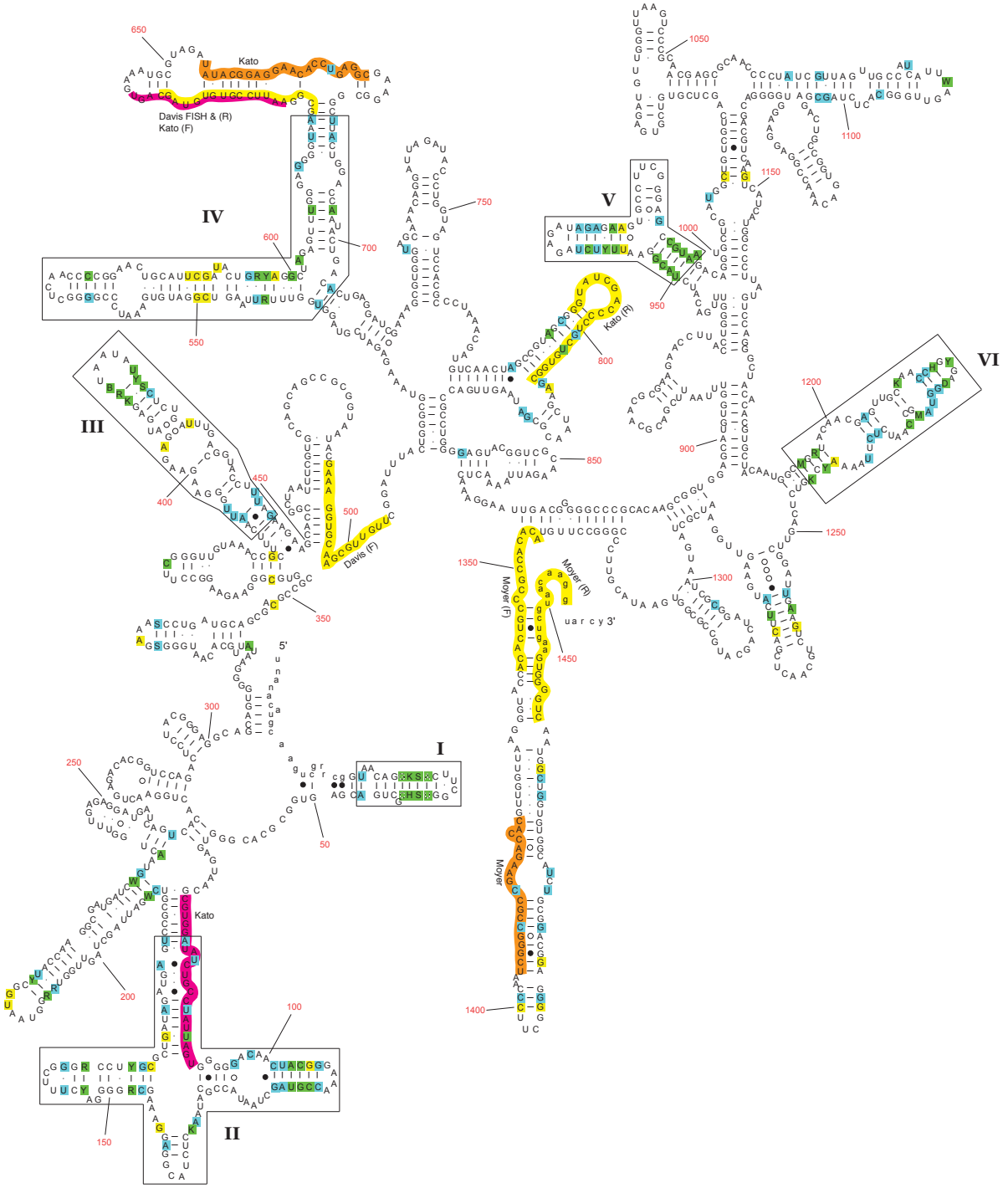


FIG. B1. Small subunit ribosomal RNA (SSU rRNA) secondary structure analysis of the consensus sequence for *Zetaproteobacteria* OTU 1. Variability between OTU 1 and the consensus sequence for OTU 2, OTU 15, or OTUs 2 and 15, is indicated by a yellow, blue, or green highlighted base, respectively. Pink, yellow, and orange highlighted runs correspond to FISH probes, Q-PCR primers, and TaqMan probes, respectively. Six regions with relatively high variability are identified.